

文章编号:1672-3961(2018)03-0034-06 DOI:10.6040/j.issn.1672-3961.0.2017.433

基于卷积神经网络的中文财经新闻分类方法



谢志峰^{1,2}, 吴佳萍¹, 马利庄^{2,3}

(1. 上海大学影视工程系, 上海 200072; 2. 上海电影特效工程技术研究中心, 上海 200072;
3. 上海交通大学计算机科学与工程系, 上海 200240)

摘要:针对目前财经领域内新闻数据杂乱无章、缺乏自动高效管理等问题,提出一种基于卷积神经网络的中文财经新闻分类方法。收集大规模财经新闻语料,通过无监督学习方法训练获得一个广义通用的财经类词向量模型,将词向量引入到卷积神经网络模型训练中实现有效分类。与传统方法相比,基于卷积神经网络的中文财经新闻分类方法网络模型结构简单,针对小样本集也能表现优异的性能,不仅能有效解决中文财经新闻分类问题,还可充分证明卷积神经网络在处理文本分类问题中的有效性。

关键词:财经新闻;卷积神经网络;文本分类;词向量;自然语言处理;深度学习

中图分类号:TP391.1 **文献标志码:**A

引用格式:谢志峰,吴佳萍,马利庄. 基于卷积神经网络的中文财经新闻分类方法[J]. 山东大学学报(工学版),2018,48(3):34-39,66.

XIE Zhifeng, WU Jiaping, MA Lizhuang. Chinese financial news classification method based on convolutional neural network[J]. Journal of Shandong University(Engineering Science), 2018, 48(3):34-39,66.

Chinese financial news classification method based on convolutional neural network

XIE Zhifeng^{1,2}, WU Jiaping¹, MA Lizhuang^{2,3}

(1. Department of Film and Television Engineering, Shanghai University, Shanghai 200072, China;
2. Shanghai Engineering Research Center of Motion Picture Special Effects, Shanghai 200072, China;
3. Department of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai 200240, China)

Abstract: In order to complete the task of financial news classification, a new method based on convolutional neural network for the classification of Chinese financial news was presented. A simple CNN was trained with one layer of convolution on top of word vectors obtained from an unsupervised neural language model. These vectors were trained on a large number of financial news corpus. Compared with the traditional methods, the network model based on convolutional neural network was simple in structure, which could show excellent performance by using small sample set. The method not only could solve the Chinese financial news classification problem effectively, but also prove the effectiveness of convolutional neural network in dealing with problems of text classification fully.

Key words: financial news; convolutional neural network; text classification; word vector; natural language processing; deep learning

0 引言

财经新闻的内容涉及与经济相关的所有领域,包括从生存到就业,从农村到城市,从经济到社会、政治生活等,简练的文字涵盖海量的信息。同时由

于互联网的出现,信息传播速度得到迅速提高,读者可以第一时间了解相关信息,特别对企业家与投资者,财经新闻是其判断市场变化的重要依据。但面对新闻数据呈爆炸式增长且杂乱无章的现象,用户获取有用信息的难度同比增加,对海量财经新闻正负性的分类存在迫切需要,然而当前对于财经新闻

收稿日期:2017-08-29; 网络首发时间:2018-05-03 17:21:33

网络首发地址:<http://kns.cnki.net/kcms/detail/37.1391.t.20180503.0847.004.html>

基金项目:国家自然科学基金资助项目(61303093,61402278,61472245);上海市科委科技攻关资助项目(16511101300)

作者简介:谢志峰(1982—),男,江苏如东人,博士,讲师,主要研究领域为信息处理、深度学习等. E-mail:zhifeng_xie@shu.edu.cn

的分类研究却很少,至今没有提出一种有效的分类方法。利用计算机技术对财经新闻进行有效分类,是一个急需解决的问题。

实现财经新闻分类的技术基础是自然语言处理(natural language processing, NLP)中的文本分类技术。应用文本分类技术可以解决信息杂乱无序的问题,较大程度地帮助用户实现信息的查找定位。目前,传统的文本分类方法^[1]及其存在的缺陷有:(1)支持向量机(support vector machine, SVM),计算复杂,不能通用于非线性问题的解决;(2)决策树分类,不适用训练大样本集,容易受噪声数据影响产生过拟合;(3)朴素贝叶斯分类,实际应用条件与假设条件不符,不满足各属性相互独立,影响正确分类。

相比传统的分类方法,近年来兴起的人工神经网络方法则表现出了良好的智能特性:(1)自适应性,对周围环境具有学习适应能力;(2)容错性,对噪声和信息缺失影响相对不敏感;(3)模式识别性,能很好地识别多变量的模式。

借助人工神经网络方法已经很好地解决了很多实际应用问题,如模式识别、智能机器人、医学、自然科学^[2]等等。近年来,人工神经网络中的卷积神经网络因具有相当强大的局部特征提取能力,已在计算机视觉、语音识别等领域取得良好的成果。经典的 AlexNet^[3]、GoogleNet^[4]、VGGNet^[5]等模型在图像识别上有良好的表现,但模型结构复杂,参数总量大,很难应用在自然语言处理中。

文本分类是自然语言处理中常见的任务,根据学习方法可分为有监督学习和无监督学习。目前大部分研究方法都采用有监督的学习方式,如基于文本特征向量相关性方法、基于遗传算法的方法、基于关联的方法、基于期望最大化算法(expectation maximization algorithm, EM)等都是典型的有监督学习方法^[6]。2003年,Bengio等^[7]最早提出用神经网络构建二元语言模型(neural network language model, NNLM),实现将词映射转换到低维实数向量,指出可以通过词之间的距离来判断各词的语义相似度。同时结合非线性神经网络提出了 n-gram 模型,随后使用随机梯度下降法优化模型,如此得到的语言模型自带平滑,性能提高 20%。2009年,Andriy Mnih 等人^[8-9]从限制玻尔兹曼机(restricted boltzmann machine, RBM)出发,逐步修改能量函数,得到更为简单的“Log-Bilinear”模型,在实际训练速度和预测速度上都有很大的提升。2011年,Socher 等人^[10]使用基于词向量的递归神经网络,实现对文本的向量表示,该网络可以有效用于文本分析。同年,Col-

lobert 等^[11]结合多层一维卷积神经网络实现了同时包含处理词性标注、语块切分、命名实体识别、语义角色标注 4 个典型自然语言处理任务的 SENNA 系统。2013 年,Mikolov^[12]借鉴 Log-Bilinear 模型的思想,通过添加隐藏层的多次递归提高了语言模型性能,在词预测的准确率方面超越了当时最好的基准系统。Zhou 等^[13]则提出一种称为主动深度网络(active deep network, ADN)的半监督学习算法用于解决情感分类问题,但网络复杂度较高。2014 年,John Son^[14]提出基于词表达方式的卷积神经网络模型,显示了词序特征在文本分类中的有效性。同年,Blunsom 等^[15]提出动态卷积神经网络句子语义模型,在 Stanford Sentiment Treebank 的二元分类与多元分类上再次取得一定的进步。2015 年, Kim^[16]将 Collobert 构建的 CNN 模型结合 1 000 亿个单词训练的向量模型应用于电影评论分类中,取分类准确率为 88.1%,这是当时最好性能。在硬件方面,基于神经网络的预训练过程通常需要高性能计算机的支持,如图形处理单元(Graphic Processing Unit, GPU)、多核计算等。同时 RBM、Auto-encoders 为深度神经网络结构中的组成单元提供算法支持。

本研究提出将卷积神经网络应用到中文财经新闻分类中,主要贡献包括以下两点:(1)通过收集大量中文财经新闻数据,结合 Word2vec 训练得到一个中文财经类词向量模型,该模型能够将财经词汇映射为多维实数向量,为接下来更高层级的训练模型提供底层支持;(2)在中文财经类词向量模型的基础上,提出一个基于卷积神经网络(convolutional neural network, CNN)的中文财经新闻分类方法,该方法具有参数少、模型简单等特点,相比于其他传统分类方法能够得到更精确的分类结果。

1 中文财经新闻分类

1.1 整体流程

本研究将 Word2vec 训练出的词向量作为候选特征向量,把文本转换成一个由词序列和词向量嵌套而成的二维矩阵,再通过 CNN 对其进行处理。主要分为词向量训练、文本预处理、神经网络模型训练、新闻分类四个部分,如图 1 所示。

(1)词向量训练。将分好词的语料文本使用神经网络来训练,在训练过程中生成一组向量,将每个词映射到一个 n 维词向量。

(2)文本预处理。利用训练好的词向量模型将

待训练样本转化为以词向量为元素的序列表示,得到一个 2 维矩阵,用于后续神经网络的训练。

(3)CNN 模型训练。将训练集通过多块卷积神

经网络进行有监督的学习训练。

(4)新闻分类。通过使用 softmax 方式完成对句子的分类。

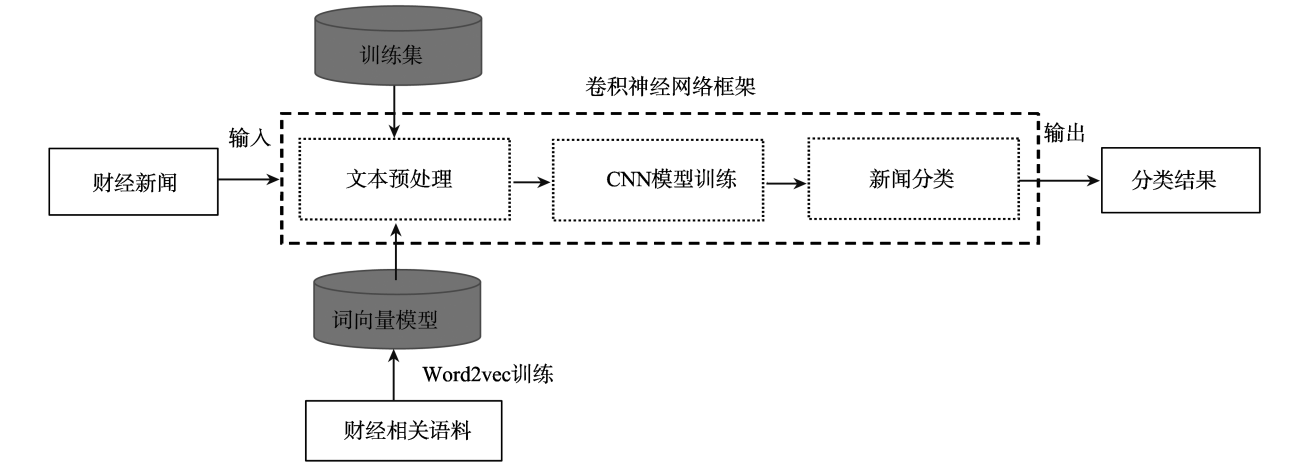


图 1 整体流程
Fig.1 The overall process of the method

1.2 词向量训练

词的向量化就是将语言中的词进行数学化表示。Mikolov 等^[17]通过借鉴 Bengio^[7]提出的 NNLM 以及 Hinton^[18]的 Log_Linear 模型在 2013 年提出了连续分布式词向量表示模型(Word2vec 语言模型)。本研究使用 Google 公司开源的 Word2vec 工具训练了一个财经类词向量模型。

1.2.1 语料采集

本研究采用的词向量训练语料来源包含三部分:(1)新浪新闻。新浪财经作为当下最具影响力的主流媒体平台,财经领域内重要会议及事件的报道达 80%。本研究共采集新浪财经“7×24 小时全球实时财经新闻”平台上 28 万条最新财经资讯。(2)财经词汇百度百科。采集搜狗词库 12 965 个财经相关词汇的百度百科内容。(3)股票简介。采集同花顺财经网上证 A 股、深证 A 股、中小板及创业板全部共 2 768 个股票相关简介及百度百科内容。

1.2.2 数据处理

对数据的处理包括去除非法字符、停用词处理、分词处理、编码处理四部分。中文文本与英文文本主要不同之处是不存在天然分隔符,因此需要进行分词处理。同时,中文文本都存在半角、全角两种字符,需要对文本进行编码规范。

1.2.3 Word2vec 训练

Word2vec 模型有两种,分别是连续词袋(continuous bag of words, CBOW)模型及 Skip-gram 模型^[19]。本研究采用的是 Skip-gram 模型。训练参数如表 1 所示,训练语料中共包含 55 745 009 个词汇,

通过训练共得到 222 950 个词汇的向量表示。以“中国银行”为例,可以得到其向量位置及其与其他词汇向量的相似度,见图 2。

表 1 Word2vec 训练参数					
Table 1 Training parameters of Word2vec					
模型	向量 维度	窗口 大小	词汇 数量	词向量 数量	训练 时间/min
Skip-Gram	300	8	55 745 009	222 950	1 461



图 2 Word2vec 中词语相似度比较
Fig.2 Similarity comparison of Word2vec vocabulary

训练所得的财经类词向量可用于完成财经类自然语言处理 NLP 里面的各种任务,如词性标注、命名实体识别、短语识别、语义角色标注等。

1.3 卷积神经网络

1.3.1 CNN 模型结构

卷积神经网络是前馈神经网络的一种,由卷积层和池化层叠加组合而成。卷积层作为特征提取层,通过滤波器提取局部特征,经过卷积核函数运算产生特征图,输出到池化层^[20]。池化层属于特征映射层,对卷积层产生的特征图进行降采样,输出局部最优特征。图 3 为本研究采用的卷积神经网络模型。

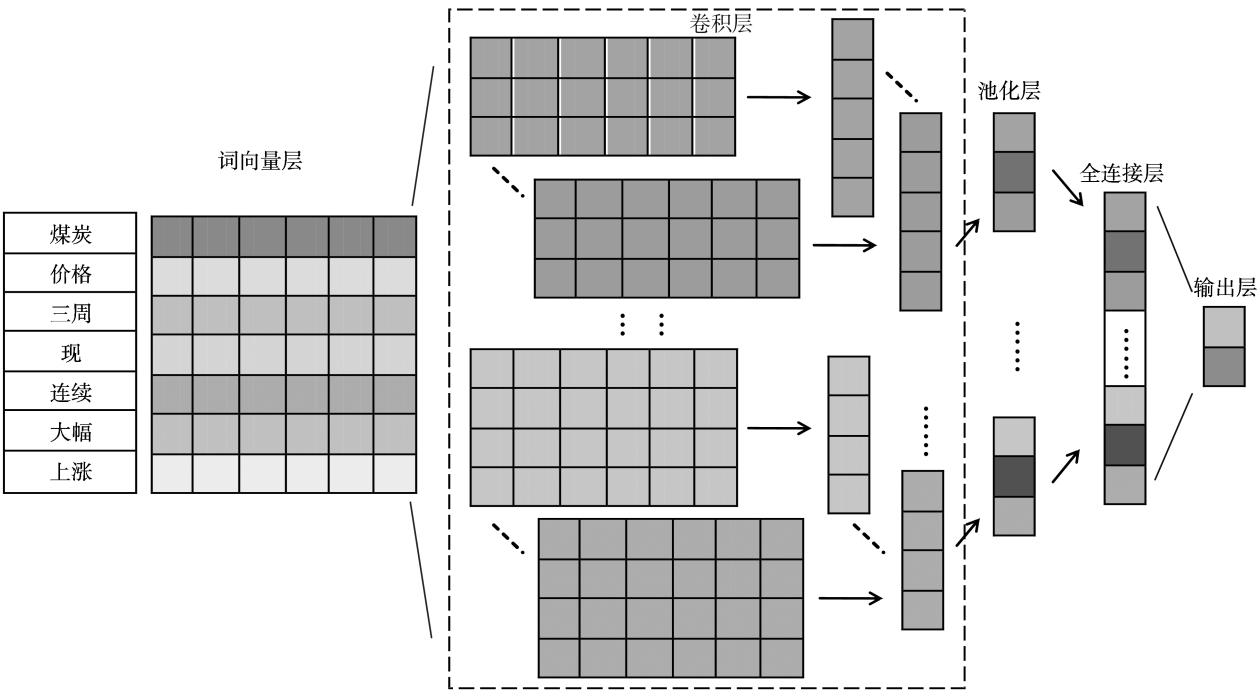


图3 卷积神经网络模型
Fig.3 Convolution neural network model

(1) 词向量层。将词序列中的元素转化为词向量,令 $\mathbf{x}_i \in \mathbf{R}^k$ 为第 i 个词对应的 k 维词向量,长度为 n 的句子表示为

$$\mathbf{x}_{1:n} = \mathbf{x}_1 \oplus \mathbf{x}_2 \oplus \mathbf{x}_3 \oplus \cdots \oplus \mathbf{x}_n, \tag{1}$$

式中: \oplus 为级联运算符,则 $\mathbf{x}_{i:i+h-1}$ 可以表示为由第 i 到第 $i+h-1$ 个词组成的局部特征矩阵。

(2) 卷积层。对不同尺寸的滤波器建立不同的卷积层,保持滤波器维度与词向量维度一致,若滤波器窗口大小为 h ,则对输入的特征矩阵进行卷积操作后,生成特征

$$\mathbf{c}_i = f(\mathbf{w} \cdot \mathbf{x}_{i:i+h-1} + \mathbf{b}), \tag{2}$$

式中: $\mathbf{w} \in \mathbf{R}^{h \times k}$ 为滤波器的权重矩阵; $\mathbf{b} \in \mathbf{R}$ 为偏差项; f 是卷积核函数。该滤波器应用于句子 $\{\mathbf{x}_{1:h}, \mathbf{x}_{2:h+1}, \cdots, \mathbf{x}_{n-h+1:n}\}$ 得到特征图

$$\mathbf{C} = [\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4, \cdots, \mathbf{c}_{n-h+1}]. \tag{3}$$

(3) 池化层。池化层采用 Collobert^[12] 提出的 max-over-time pooling 方法进行特征采样,对卷积层的输出特征进一步抽象,保留最重要的特征:

$$\hat{\mathbf{C}} = \max \{\mathbf{C}\}. \tag{4}$$

(4) 全连接层。全连接层由池化层输出的多个特征向量构成,若有不同尺寸的滤波器 p 种,每种滤波的数量为 q ,则全连接层的特征向量

$$\mathbf{V} = [\hat{\mathbf{C}}_{1,1}, \cdots, \hat{\mathbf{C}}_{1,q}, \hat{\mathbf{C}}_{2,1}, \cdots, \hat{\mathbf{C}}_{2,q}, \cdots, \hat{\mathbf{C}}_{p,1}, \cdots, \hat{\mathbf{C}}_{p,q}]. \tag{5}$$

(5) 输出层。使用全连接的 softmax 层完成分类任务。

1.3.2 数据预处理

目前公开的新闻数据集少有中文财经类数据,本研究采集的数据集共包含 8 138 条财经新闻,人工标注为正负两类,对经济市场产生积极正面影响的标注为正,对经济市场产生滞后影响或报道负面的标注为负,如表 2 所示。经标注,含正向新闻 4 121 条,负向新闻 4 017,最长文本长度为 50。

表2 人工标注示例	
Table 2 The sample of manual annotation	
新闻文本	人工标注
德国总理默克尔:德国经济表现良好。	正
欧洲 Stoxx600 指数开盘上涨 0.2%,报 344.77。	正
伊朗国家石油公司将签署价值达 25 亿美元的石油合约。	正
瑞银:低利率催生楼市新泡沫。	负
中国工商银行纸原油周四亚市早盘震荡下跌。	负
南方黑芝麻糊年内两次上黑榜,3 批次产品不合格。	负

本研究利用 JIEBA 分词工具对试验数据集进行分词、去除非法字符、停用词处理。词向量采用上文已经训练完成的 Word2vec 向量模型,共包含 22 万个汇,词汇覆盖率为 87.36%。

1.3.3 超参数和训练

对于卷积神经网络模型,本研究采用表 3 所示的卷积神经网络模型可调参数设置,使用窗口大小为 3、4、5 的滤波器各 100 个。

在模型训练阶段,采用 Zeiler^[21] 在 2012 年提出的 Adadelta Update Rule 进行随机梯度下降^[22] 更新

模型,参数取 0.5。

表 3 试验模型参数
Table 3 Parameters of the CNN model

卷积核函数	词向量维度	滤波器窗口大小	过滤器数量	随机更新参数比例	训练迭代次数
rectified linear	300	3、4、5	100	0.5	25

2 试验结果

本研究使用静态词向量和动态词向量两种不同形式的通道模式。(1)CNN 静态词向量。使用 Word2vec 词向量,在训练过程中保持所有词向量静态。若存在新词汇不在 Word2vec 词向量模型中,则随机初始化。(2)CNN 动态词向量。使用 Word2vec 词向量,但将其作为卷积神经网络中一项参数进行迭代训练,实现词向量的动态调整。

为了降低其他随机因素对上述变化的影响,进行十折交叉验证(10-fold cross-validation)来评估方法的准确性。将样本数据分成 10 组,轮流将 9 组数据作为训练集用于训练,剩下 1 组为测试集,即每组数据依次作为测试集使用。每次测试均得到一个正确率,最后取 10 次测试结果的平均值用于总体评估。在两种模式下的试验结果如表 4 所示。

表 4 静态与动态模式下训练结果
Table 4 Results of static and dynamic model

静态		动态	
训练时间/s	准确率/%	训练时间/s	准确率/%
1 759.65	73.60	2 986.25	76.27
1 677.52	79.42	2 988.25	79.80
1 678.60	76.21	2 991.75	77.84
1 663.44	78.76	3 001.75	80.46
1 678.54	78.31	2 990.06	82.60
1 666.25	74.63	2 995.25	77.62
1 694.25	76.04	2 988.54	79.60
1 736.50	77.60	2 986.27	82.16
1 834.75	79.85	2 998.65	82.93
1 817.50	75.05	2 982.29	81.24
1 720.70	76.96	2 990.91	80.06

在静态模式下,正确率分布为 73.69%~79.85%,平均正确率为 76.96%。

在动态模式下,正确率分布为 76.27%~82.93%,平均正确率为 80.06%。

相比于静态模式,动态模式虽然在训练速度上没有很好的表现,但通过对词向量的动态调整,其分类性能优于静态模式,提高 3%,可以有效提升模型的分类能力。

结果表明,两种模式都能完成对大部分财经新闻的准确分类:如样例“德国总理默克尔:德国经济表现

良好”“欧洲 Stoxx600 指数开盘上涨 0.2%,报 344.77。”“中国工商银行纸原油周四亚市早盘震荡下跌”包含“良好”“上涨”“下跌”等明显表示正负性词且结构较简单的句子都能做出正确的分类。但是,如“LendingClub 今日股价下跌 8.2%,本周累计上涨 9.1%”这一样例同时包含“下跌”“上涨”正负两类词性,前后语义发生转折,此类样本在训练集中出现较少,分类结果存在一定错误率。

因财经新闻涉及领域较广,覆盖与经济相关的全部内容,其对词向量的覆盖率要求较高,且相比于英文文本,中文文本需要进行更多的数据处理。本研究方法已经在基于小训练集的基础上取得了良好的结果。为进一步显示方法的有效性,本研究特引入基于深度学习的方法及朴素贝叶斯 SVM 等 5 种机器学习方法,使用相同的试验数据作为输入,试验结果如表 5 所示。表 5 中:CNN_S 为静态卷积神经网络方法,CNN_US 为动态卷积神经网络方法,MVRNN(matrix-vector recurrent neural networks)为矩阵-向量递归神经网络方法^[23],Sent-Parser 为情感分析特定解析器方法^[1],NBSVM(naive bayes support vector machine)为朴素贝叶斯 SVM 方法^[24],MNB(multi-element naive bayes)为多元素朴素贝叶方法^[24],Tree-CRF(Tree-conditional random field)为结合条件随机依赖树的方法^[25]。

表 5 各分类方法性能对比
Table 5 Performance comparison of classification methods

模型	识别率/%	模型	识别率/%
CNN_S	80.06	NBSVM	74.78
CNN_US	76.96	MNB	74.41
MVRNN	74.57	Tree-CRF	72.42
Sent-Parser	75.51		

结果显示,本研究方法分类性能比其余 5 种方法相比更具竞争力。基于卷积神经网络方法在两种模式下准确率均高于其他方法,尤其在动态模式下,准确率比其余最高的 Sent-Parser 方法高 4.55%,比最低的 Tree-CRF 方法高 7.64%。

3 结论

本研究提出了一种基于卷积神经网络的中文财经新闻分类方法。首先,收集大量新闻语料,利用词

的上下文信息训练得到一个财经类词向量模型,实现将财经词汇映射为低维实数向量。其次,将词向量组合成二维矩阵作为训练样本的深度特征,引入到卷积神经网络模型训练中。该模型充分利用词语本身的特征及上下文语义信息,将整条新闻映射成一个整体向量作为模型输入,在保证分类性能的同时简化模型复杂度。

试验结果证明,本研究使用的模型结构简单,训练速度快,相比传统分类方法,分类准确率有大幅提升,可以有效解决中文财经新闻的分类问题。而通过 Word2vec 训练得到的财经类词向量作为广义通用的语言信息,可重复用作不同神经网络的句子输入,同时也表明 Word2vec 在自然语言处理中具有重要作用。

参考文献:

- [1] WANG S, MANNING C D. Baselines and bigrams: Simple, good sentiment and topic classification[C]//Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics. Jeju Island, Korea: ACL, 2012: 90-94.
- [2] KRIZHEVSKY A, SUTSKEVER I, HINTON G. Imagenet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2012, 60(2): 1097-1105.
- [3] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]//Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA: IEEE, 2015: 1-9.
- [4] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. [2017-06-20]. <https://arxiv.org/pdf/1409.1556v6.pdf>.
- [5] 徐姗姗,刘应安,徐昇.基于卷积神经网络的木材缺陷识别[J].山东大学学报(工学版),2013,43(2):23-28.
XU Shanshan, LIU Ying'an, XU Sheng. Wood defects recognition based on the convolutional neural network[J]. Journal of Shandong University (Engineering Science), 2013, 43(2): 23-28.
- [6] 奚雪峰,周国栋.面向自然语言处理的深度学习研究[J].自动化学报,2016,42(10):1445-1565.
XI Xuefeng, ZHOU Guodong. A survey on deep learning for natural language processing[J]. Acta Automatica Sinica, 2016, 42(10): 1445-1565.
- [7] BENGIO Y, DUCHARME R, VINCENT P, et al. A neural probabilistic language model[J]. The Journal of Machine Learning Research, 2006, 3(6):1137-1155.
- [8] FISCHER A, IGEL C. An introduction to restricted Boltzmann machines [C]//Iberoamerican Congress on Pattern Recognition. Berlin, Germany: Springer, 2012: 14-36.
- [9] MNH A, HINTON G. A scalable hierarchical distributed language model [C]//International Conference on Neural Information Processing Systems. Vancouver, Canada: NIPS, 2008: 1081-1088.
- [10] SOCHER R, PENNINGTON J, HUANG E H, et al. Semi-supervised recursive autoencoders for predicting sentiment distributions[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Edinburgh, Britain: ACL, 2011: 151-161.
- [11] COLLOBERT R, WESTON J, BOTTOU L, et al. Natural language processing (almost) from scratch[J]. The Journal of Machine Learning Research, 2011, 12(1): 2493-2537.
- [12] MIOLOV T, SUTSKEVER I, CHEN K, et al. Distributed representations of words and phrases and their compositionality[C]//Proceedings of the 26th International Conference on Neural Information Processing Systems. Lake Tahoe, USA: NIPS, 2013: 3111-3119.
- [13] ZHOU Shusen, CHEN Qingcai, WANG Xiaolong. Convolutional active deep learning method for semi-supervised sentiment classification [J]. Neuro computing, 2013, 120(10): 536-546.
- [14] JOHNSON R, ZHANG T. Effective use of word order for text categorization with convolutional neural networks[J]. Eprint Arxiv, 2014:1412.
- [15] BLUNSOM P, GREFFENSTEEN E, KALCHBRENNER N. et al. A convolutional neural network for modelling sentences[C]//Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics. Baltimore, USA: ACL, 2014: 655-665.
- [16] KIM Y. Convolutional neural networks for sentence classification[C]//Proceedings of the EMNLP. Doha, Qatar: Association for Computational Linguistics, 2014: 1746-1751.
- [17] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space [EB/OL]. [2017-06-20]. <https://arxiv.org/pdf/1301.3781v3.pdf>
- [18] HINTON G E, SRIVASTAVA N, KRIZHEVSKY A, et al. Improving neural networks by preventing co-adaptation of feature detectors[EB/OL]. [2017-06-20]. <https://arxiv.org/pdf/1207.0580v1.pdf>.
- [19] 唐明,朱磊,邹显春.基于 Word2vec 的一种文档向量表示[J].计算机科学,2016,43(6):264-269.
TANG Ming, ZHU Lei, ZOU Xianchun, et al. Document vector representation based on Word2vec[J]. Computer Science, 2016, 43(6): 264-269.

- [10] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]//International Conference on Neural Information Processing Systems. Nevada, USA: Curran Associates Inc, 2012:1097-1105.
- [11] GUO L, LI F, LIEW W C. Image Aesthetic Evaluation Using Parallel Deep Convolution Neural Network[C]//International Conference on Digital Image Computing: Techniques and Applications. Gold Coast, Australia: IEEE Computer Society, 2016:1-5.
- [12] LU X, LIN Z, JIN H, et al. Rapid: rating pictorial aesthetics using deep learning [C]//Proceedings of the ACM International Conference on Multimedia. Orlando, USA: ACM, 2014:457-466.
- [13] WANG L, OUYANG W, WANG X, et al. Visual tracking with fully convolutional networks [C]//IEEE International Conference on Computer Vision. Santiago, Chile: IEEE Computer Society, 2015:3119-3127.
- [14] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: IEEE Computer Society, 2014:580-587.
- [15] SUN Y, WANG X, TANG X. Deep learning face representation from predicting 10 000 classes[C]//IEEE Conference on Computer Vision and Pattern Recognition. Hong Kong, China: IEEE, 2014:1891-1898.
- [16] KARAYEV S, TRENTACOSTE M, HAN H, et al. Recognizing image style[EB/OL]. (2014-7-23) [2018-3-22]. <https://arxiv.org/abs/1311.3715>.
- [17] CHATFIELD K, SIMONYAN K, VEDALDI A, et al. Return of the devil in the details: delving deep into convolutional nets [C]//Proceeding of British Machine Vision Conference. Nottingham, UK: BMVA Press, 2014.
- [18] JIA Y, SHELHAMER E, DONAHUE J, et al. Caffe: convolutional architecture for fast feature embedding [EB/OL]. (2014-6-20) [2018-3-22]. <https://arxiv.org/abs/1408.5093>

(编辑:甄鹏)

(上接第 39 页)

- [20] 陈钊,徐睿峰,桂林,等. 结合卷积神经网络和词语情感序列特征的中文情感分析[J]. 中文信息学报, 2015, 29(6): 172-178.
- CHEN Zhao, XU Ruifeng, GUI Lin, et al. Combining convolutional neural networks and word sentiment sequence features for Chinese text sentiment analysis[J]. Processing of Journal of chinese Information, 2015, 29(6): 172-178.
- [21] ZEILER M. ADADELTA: An adaptive learning rate method[EB/OL]. [2017-06-20]. <https://arxiv.org/pdf/1212.5701v1.pdf>.
- [22] BOTTOU L. Large-scale machine learning with stochastic gradient descent[C]//Proceedings of COMPATAT' 2010. Berlin, Germany: Springer, 2010: 177-186.
- [23] SOCHER B, HUVAL C, MANNING A. Semantic compositionality through recursive matrix-vector spaces [C]//Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. Jeju Island, Korea: ACL, 2012: 1201-1211.
- [24] DONG L, LIU S J, ZHOU M, et al. A statistical parsing framework for sentiment classification[J]. Computational Linguistic, 2015, 41(2): 293-336.
- [25] NAKAGAWA K, INUI S, KUROHASHI S. Dependency tree-based sentiment classification using CRFs with hidden variables[C]//The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics. Los Angeles, USA: ACL, 2010: 786-794.

(编辑:陈燕)