

文章编号: 1003-0077(2018)12-0118-07

基于自联想记忆与卷积神经网络的跨语言情感分类

刘 娇, 崔荣一, 赵亚慧

(延边大学 计算机科学与技术学院 智能信息处理研究室, 吉林 延吉 133000)

摘 要: 该文提出了一种以商品评论为对象的基于语义融合的跨语言情感分类算法。该算法首先从短文本语义表示的角度出发, 基于开源工具 Word2Vec 预先生成词嵌入向量来获得不同语言下的信息表示; 其次, 根据不同语种之间的词向量的统计关联性提出使用自联想记忆关系来融合提取跨语言文档语义; 然后利用卷积神经网络的局部感知性和权值共享理论, 融合自联想记忆模型下的复杂语义表达, 从而获得不同长度的短语融合特征。深度神经网络将能够学习到任意语种语义的高层特征致密组合, 并且输出分类预测。为了验证算法的有效性, 将该模型与最新几种模型方法的实验结果进行了对比。实验结果表明, 此模型适用于跨语言情感语料正负面情感分类, 实验效果明显优于现有的其他算法。

关键词: 跨语言情感分类; 自联想记忆; 词共现; 卷积神经网络

中图分类号: TP391

文献标识码: A

Cross-lingual Sentiment Classification Based on Auto-associative Memory and Convolutional Neural Networks

LIU Jiao, CUI Rongyi, ZHAO Yahui

(Intelligent Information Processing Lab., Department of Computer Science and Technology,
Yanbian University, Yanji, Jilin 133000, China)

Abstract: A cross-linguistic sentiment classification algorithm based on semantic fusion is proposed for product reviews. First, information of different languages is generated by the open-source tool Word2Vec in advance. Then, the auto-associative memory relationship is proposed to extract the cross-lingual document semantic, according to statistical relevance of word vector between different languages. Local perception and weight sharing techniques of convolutional neural networks are applied to amalgamate of complex semantic expression in auto-associative memory model, so as to generate the phrase features of different lengths. The dense combination of high-level semantic features is learned by deep neural network for all languages, which paves the way for classification predictions. It is demonstrated that, for positive and negative sentiment classification of cross-lingual sentiment corpus, the proposed model is much more effective than other existing algorithms

Keywords: cross-lingual sentiment classification; auto-associative memory relationship; word co-occurrence; convolutional neural networks

0 引言

情感分类属于较为典型的二分类问题, 即给含有情感色彩的文档一个态度偏向, 支持或者反对。西方语言对情感分类研究起步较早, 具有丰富的情

感词典语料等资源, 而中文情感资源相对匮乏。研究跨语言情感分类不仅是为了消除语言之间的应用屏障, 还可以将资源丰富型语言的研究资源应用到资源匮乏型语言中去, 帮助其他语言发展, 跨越语言之间的鸿沟。本文提出的自联想记忆模型可以减小资源不均衡对分类精度带来的影响, 适用于跨语言

收稿日期: 2018-07-11 定稿日期: 2018-08-24

基金项目: 国家语委“十三五”科研规划项目(YB135-76); 延边大学外国语言文学世界一流学科建设科研项目(18YLPY13)

情感分类任务。

跨语言情感分类领域近年来显现出许多优秀的模型。文献[1]提出混合自训练模型和协同训练模型,分别训练两个语种的数据,采用双视图来得到最终的分类结果,分类性能超过了自训练和协同训练中最好的结果,并在当时取得了 NLPCC 跨语言情感分类评测任务(2013)的第一名。文献[2]在解决跨语言模型的问题时采用迁移学习的方法,将源语言和目标语言分别视作两个领域,在此引入基于样例迁移的监测机制来避免发生负迁移情况,可以减少样本数据中的噪音,提高分类性能。文献[3]结合双视图与迁移学习两种方法构建模型。双视图中将英文视为标准的参考知识,并在英文空间中进行训练,通过对英文样本的翻译来得到中文视图中的数据,并引入知识验证函数来过滤中文翻译样例中由于翻译产生的可能存在的噪音。

深度学习模型在情感分类上的应用研究也取得了不小的进展。文献[4]通过注意力机制来捕获不同上下文信息对给定特征的重要性,将注意力机制与 LSTM 结合起来对句子进行语义建模。文献[5]提出一个端到端的对抗记忆网络去解决跨领域的情感分类。使用注意力机制来自动抓取特征,模型框架包括两个参数共享的记忆网络。其中,一个用于情感分类,一个用于主题分类,联合训练的同时自动选择特征以最小化情感分类的误差。文献[6]提出了使用降噪自编码器或者长短时记忆递归网络(LSTM)结合上下文情感模糊词,以学习中英文双语特征的信息表示。

在跨语言语义表示方面,融合和转换语言空间的方法也是研究的重点。Gliozzo 等^[7]通过使用比较语料库对英语和意大利语进行跨语言文本分类,对比较语料库做潜在语义分析,建立一个由两种语言的潜在语义构成的低维投影空间,并将两种语言构成的词—文档矩阵都投影到该空间以后进行分类。Hanneman 等^[8]通过构造基于句法的全文翻译算法来提高分类的正确率。Faruqui 等利用典型相关性分析^[9-10]进行跨语言文本分析,即找出两个语言空间下相关系数最大的向量,建立起跨语言桥梁。而本文以翻译语料库为基础获得评论语料在不同语种下的表示,利用不同语种的词汇之间的共现程度和它们对应词嵌入的空间关系来构造跨语言自联想记忆关系。在合并的语言空间下,使用扩展的卷积神经网络模型进行跨语言情感分类任务。证明了在对跨语言评论语料进行情感正负面分类时,语义的

融合方法比转换方法更有效,即使在小规模数据集下也可以获得较高的分类精度。

1 跨语言自联想记忆构造算法

自联想记忆是指模型能够学习并回忆不同对象个体间的关系。比如,精通中英双语的人通过中文文本可以联想到同一语义的英文文本,任意给出一种语言的信息都可进行联想以获得该文本语义在融合语种模型空间中的完整表示。由文献[11]中的分布假设推导可知单词的关联程度可以用基于语料库的方法来获得。在一个大规模语料库中,每个文档中字词的分布都可以用向量来表示,词与词之间的关联程度也可以用这个向量来计算。通常在平行语料库中,如果分别属于两种语言的两个词出现在同一篇语义的文档中,一般情况下我们可以确定这两个词汇具有高度的语义相关性。根据这个关系,对特定语言中的每一个词汇,找出另一种语言中与该词汇相关度最大的词,可以得到一种语言关于另一种语言的共现词汇表。互相共现的两个不同语种的词不一定能互译,但在语义上具有高度相关性^[12],它们之间的关系已经被证明适用于跨语种文档检索和相似度计算。

1.1 共现词汇表

根据 Word2Vec 模型的计算原理可知^[13],它计算词向量时体现的是词和它的局部上下文的共现关系。根据上下文窗口的滑动来迭代计算所有词的向量表示,词之间具有高度的可比性,词之间的语义距离可以在向量的空间中衡量,这证明了词之间可以直接用向量线性表示^[14]。在不同的语言中,即使源语言文档 d_s 和目标语言文档 d_t 存在互译关系,文档中的词在各自语种的语料库中分布相近,但不具有上下文关系。即 d_s 和 d_t 中的词不能在同一个词上下文中窗口中计算得到。因此,它们只能遵循如“ $v_{\text{France}} - v_{\text{法国}} \approx v_{\text{Italy}} - v_{\text{意大利}}$ ”这样的语义关系。本文结合词的共现计算方法和词嵌入向量的特点设计了共现词对如式(1)所示。

$$L_{s|t} = \{ \langle x, y \rangle \mid x \in V_s \wedge y \in V_t \wedge y = T\text{-index}(x) \} \quad (1)$$

其中:

$$T\text{-index}(j) = \arg \max_{i \in V_t} (v_i^T v_j + \alpha e^{m_{ij}}), j \in V_s \quad (2)$$

式中 V_s 和 V_t 分别表示源语言和目标语言下语

料中的词项序号集。 v_i 和 v_j 分别表示目标语言中和源语言中序号为 i, j 的词的词向量, α 代表经验参数, m_{ij} 表示这两个词共同在平行语料中的出现次数。

式(2)旨在找出两种语言中统计关系最大的两个词。在平行语料中共同出现次数最多的词 i 和词 j 在跨语言语义中具有理应有更强的关联性,但词 j 关联度最高的候选词可能不止一个。因此,公式中加入词嵌入向量 v_i 和 v_j 的内积结果来加强相关性的计算,且使共现词对的结果具有唯一性。

1.2 跨语言自联想记忆映射方法

自联想记忆是指两个同形态的数据相互之间存在联系的形式或概念构成知识存储在记忆中的具体形态。依据这个概念提出将共现词表作为两种语言的数据之间的桥梁,将自联想记忆方法应用到神经网络中。自联想记忆映射如式(3)、式(4)所示。

$$f: R^{v_s \times n} \rightarrow R^{v_t \times n} \quad (3)$$

$$f(v_{j|s}) = v_{T-index(j)|t} \quad (4)$$

式(4)中,根据输入的源语言向量可以联想到目标语言的相关向量。在跨语言任务中,只需要对不同的语言之间互相建立共现词表,输入任何语言的向量都可以联想到任意其他语言。分类模型的底层输入通常为词向量,对给定某个语种文本中的每一个词都生成一个词向量。根据自联想记忆关系,该词在任一其他语言下都有语义对应的词,依据式(3)对输入的源语言中每一个样本数据进行扩展,如图1所示。

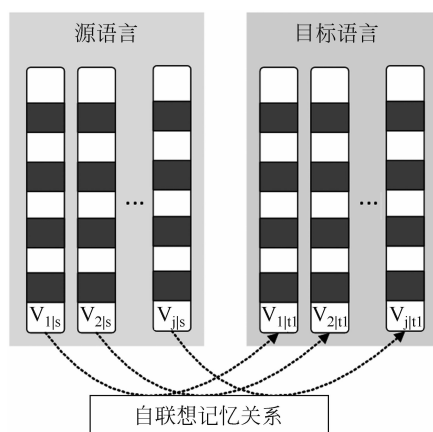


图1 基于自联想记忆的语言扩展

在 Word2Vec 的计算下,每一个词都被表示成为一个维数固定为 k 的向量。每篇文档由它的词生成的词向量以出现次序排列成一个输入矩阵,文本

信息转化为可计算的向量组合。图右侧第一个框为源语言文本的输出词向量,根据共现词汇表找到每一个词对应的目标语言下的语义关联词,在目标语言空间中的同样位置输入获得的关联词的词向量。通过自联想记忆关系,源语言中的文档矩阵 $d_{i|s}$ 可以映射到目标语言中,其文档表示为 $d_{i|t}$, $d_{i|s}$ 是 s 语言空间中 m 行 k 列的矩阵, $d_{i|t}$ 是另一种 t 语言中同样 m 行 k 列的矩阵。文本模型是将源语言空间中的矩阵拼接映射结果矩阵,如式(5)所示。

$$d = \begin{bmatrix} d_{i|s} \\ d_{i|t} \end{bmatrix}, d_{i|s}, d_{i|t} \in R^{m \times k}, d \in R^{2m \times k} \quad (5)$$

通过式(5)合并一篇文档在不同语言下的矩阵表示后,可以在合并的空间中对同语义不同语种的文档进行分类。

2 基于自联想记忆的卷积神经网络分类模型

假设自联想记忆关系产生的数据互相具有语义补充关系,因此自联想记忆关系适用于融合不同语言的语义信息。在情感分类任务中,文本的情感倾向通常只由少数词语决定,而卷积神经网络可用于提取显著特征并忽略对分类器作用较小的信息^[15]。因此本文提出基于自联想记忆卷积神经网络 MeCL-CNN (Auto-associative Memory Cross-Linguistic Convolutional Neural Network) 的跨语言情感分类方法。通过自联想记忆映射方法,文本模型可以获取任意语种文本作为输入。根据输入语种的语义映射向量作为记忆来帮助模型生成融合的语种空间,在合并的空间中进行卷积局部特征提取。

本文中卷积神经网络模型如图2所示,其中扩展的模型一共包括9层,叠加了更深的卷积层和归一化层(batch normalization)。输入层是由词嵌入向量拼接构成的文本矩阵,卷积层中每一个卷积核窗口宽度 k 与词向量宽度一致,卷积结果向量 s 的第 i 个值的计算,如式(6)所示。

$$s_i = W \cdot A_{i:i+k-1} + b, W \in R^{h \times k} \quad (6)$$

其中, A 代表输入数据, W 表示与 A 计算内积的卷积核权重参数, b 是偏置项。长度为2代表卷积核可以提取两个词组成的短语信息。同理,其他长度的卷积核也可表示对相应词数组成的短语特征的提取。卷积核的深度表示同一卷积层在相同卷积核尺寸下,卷积层由不同的特征提取规则集构成。

归一化层的计算过程,如式(7)~式(10)所示。

$$\mu_{\beta} = \frac{1}{m} \sum_{i=1}^m x_i \quad (7)$$

$$\sigma_{\beta}^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu_{\beta})^2 \quad (8)$$

$$a_i = \frac{x_i - \mu_{\beta}}{\sqrt{\sigma_{\beta}^2 + \epsilon}} \quad (9)$$

$$y_i = \gamma a_i + \beta \quad (10)$$

其中, μ_{β} 是对输入的这批数据的平均值, σ_{β}^2 是输入数据的方差, m 为输入数据的数量, 通过式(7)~式(10)来归一化数据的分布^[16], 使之具有固定的均值和方差, 由于神经网络本质上是

一个参数化模型, 不同的数据分布理应在不同的参数模型下才能更好地拟合。对于评论类型长度的情感分类语料, 当训练数据和测试数据的分布差距较大时, 模型的效果将大打折扣。此外, 在网络层数增加的过程中, 较低层网络参数的变化对高层网络的影响将越来越大, 而在式(9)的影响下, 每层的输入数据都转变成均值 0 与方差为 1 的分布。因此, 对数据的归一化可以改善这个问题, 但是在这个操作下, 网络的表达能力会变弱, 因此采用式(10)对归一化后的数据进行缩放和平移。

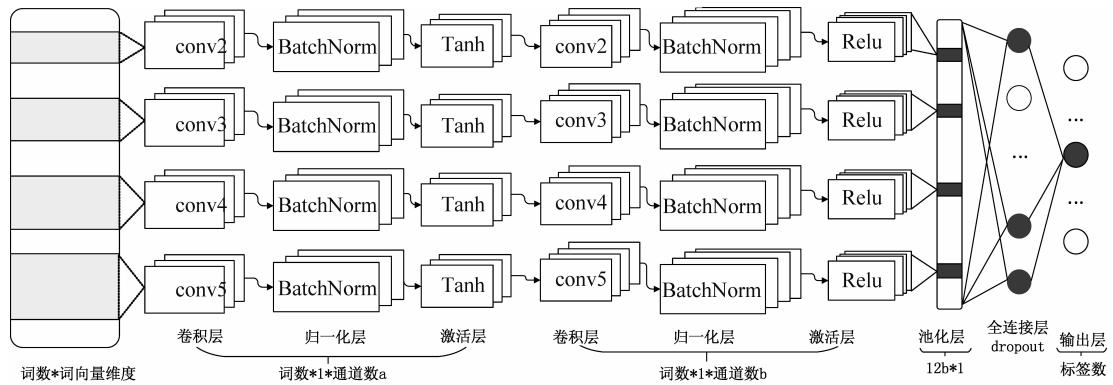


图2 扩展的 TextCNN 模型

激活层函数选择 f_{\tanh} 和 f_{relu} , 如式(11~12)所示。

$$f_{\tanh}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (11)$$

$$f_{\text{relu}}(x) = \begin{cases} x & (if \ x > 0) \\ 0 & (if \ x < 0) \end{cases} \quad (12)$$

$f_{\tanh}(x)$ 的输出区间为 $(-1, 1)$, 并且在输入数据 x 绝对值非常大时函数值达到饱和, 这使得它的作用接近生物神经元, 可抑制或刺激神经元的信息传递。因它具有均值为 0 的特性, 故收敛速度较快。relu 函数的存在极大地解决了饱和函数所遇到的问题。即在函数值饱和到 0 或 1 时, 网络层导数接近 0。因此, 会影响梯度的反向传递。在高层网络中适宜用 relu 函数来保证梯度的传输, 缓解梯度消失的问题。

池化层采用 Chunk-Max Pooling 方法。其思想是把 pooling 层中每一个深度上的向量进行等长分段, 将特征向量切割成若干段后, 只保留每个子段中最显著的特征值。在情感分类任务中, 池化层可以有效地过滤噪音, 增强分类器性能。

最后一层为分类结果输出层, 将全连接层与 softmax 层组合, 可组成 softmax 回归分类器。假设卷积层、激活层和池化层等可将向量从原始输入

映射到隐层特征空间中, 则全连接层的作用是将隐层空间中的分布式特征向量表示映射到样本标签上, 从而完成分类任务。softmax 函数如式(13)所示。

$$p_i = \frac{e^{z_i}}{\sum_k e^{z_k}} \quad (13)$$

其中, z_i 代表前一层第 i 个单元的输出, p_i 的值代表输出层上第 i 个神经元的输出, 即分类标签属于第 i 类的概率。

模型中在全连接层添加了 dropout 机制。dropout 的工作原理是在神经网络的传递过程中, 以某个特定的概率随机放弃一些单元。类似于特征选择的作用, 只保留一部分优秀的特征传递下来, 还可以降低隐含层单元之间的联合适应性。从生物学的角度来说, 它使得复杂的整段神经元的组合联动变成比较小的神经元片段之间的联合适应。在简单的线性空间中, 只学习一个空间的特征集合足以应付当前的输入, 但是在数据分布波动较大且通常存在于非线性不连续的空间中时, 只适应一个空间中的特征集合会造成过拟合的现象, 模型对与当前数据分布差别较大的输入没有识别能力, 此时对局部

空间的特征集合进行随机采样可以增强了模型的泛化能力。

3 实验结果与分析

3.1 数据集

本文实验主要对 NLPCC2013 会议中提供的“跨语言情感分类”评测任务进行实验。该语料源自 Amazon 网站的中英文产品评论数据,训练数据中含有英文语料 2 000 条。含书籍、DVD 和音乐三种商品的评论标注语料,测试数据为同等规模的中文语料。该任务针对情感的正负类提供了一批用于计算情感分类的情感词典,另有一大部分未标注的中文语料,数量是训练数据的 10 倍,所有文档都以 XML 格式进行存储。为了获得不同语言之间的关系,需要将所有标注与未标注的中文语料通过 google 翻译成英文,通过翻译语料库来获取如第 2 节所示的跨语言自联想记忆关系。实验需要采用英文语料训练本文所示模型 AMCL-CNN,并对任务重提供的中文语料进行情感正负类。经过多次实验,训练过程中参数设置如下:

(1) 设置词向量维度为 200,卷积核组合为 $[2,3,4,5]$,卷积核深度为 64;

(2) 设置 L2 正则方法且正则系数为 0.05,深度学习模型容忍度为 10;

(3) 选择带动量的随机梯度下降法,随机梯度步长为 0.01,动量系数为 0.5。

3.2 Dropout 对情感分类的影响

dropout 率的经验值一般为 0.5,但通过实验可以发现,在情感分类任务中,dropout 率对分类的正确率影响非常大。

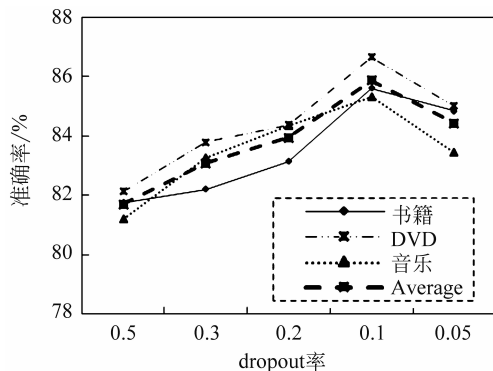


图 3 不同 dropout 的分类正确率

表 1 不同 dropout 的分类正确率

dropout	正确率/%			
	书籍	DVD	音乐	Average
0.5	81.75	82.13	81.20	81.69
0.3	82.20	83.79	83.28	83.09
0.2	83.14	84.38	84.35	83.95
0.1	85.62	86.67	85.32	85.87
0.05	84.84	85.01	83.46	84.43

为了更直观地体现实验结果的整体差别,对三种商品的评论分类的正确率计算宏平均。与上一个任务保持参数不变的情况下,不同的 dropout 参数对结果精度的影响较大,如图 3 所示。dropout 的经验值一般选取 0.5,而从图中可以看出,当 dropout 的比例取 0.5 时,各项类别的情感分类效果都为最低。当单元的保留比例增高时,分类精度也随之增高,但是在 dropout 比例为 0.1 时达到了最高。因 dropout 的本质是在神经元的信号传输过程中放弃一部分单元的工作能力,故将这些神经单元看作特征。dropout 自动完成了对输入信号进行特征选择的过程。从实验结果的角度分析,情感分类的语料特点是需要上下文特征来协助分类,又因情感分类是二分类任务,故不需要通过 dropout 来提高模型的复杂程度。这个实验证明深度学习训练过程中,不同的数据分布和任务需要不一样的超参数。

3.3 跨语言文档相似度计算结果

对跨语言情感分类研究做了对比实验。实验数据中没有平行语料,采用机器翻译获得双语语料,继而根据式(2)计算词嵌入向量之间的距离,与上一个实验过程相同,采用基于 CNN 的自联想记忆模型。实验结果如表 2 所示。

表 2 不同方法的分类正确率

算法	正确率			
	书籍	DVD	音乐	Average
文献[1]	78.70	79.65	78.30	78.89
文献[2]	80.10	81.65	78.60	80.10
文献[3]	84.65	85.18	80.93	83.59
文献[6]	82.03	81.68	79.20	80.97
本文方法	85.62	86.67	85.32	85.87

文献[1]中双视图与自训练协同训练相结合的

方法在当时取得了最好成绩。它的不足在于分别训练源语言和目标语言的分类模型时,没有充分利用两种语言之间的语义信息。文献[2]和文献[3]均基于迁移学习的方法来实现跨语言的过程继而进行情感分类。其中文献[3]取得的平均准确率高达83.59%,但它存在的明显问题是同一个模型在3个类别上的正确率差距太大。这是由于此模型中采用任务提供的未标注数据来做机器翻译的可信度认证,所以测试集的结果会在一定程度上依赖未标注数据的翻译质量和数据分布。换言之,当未标注数据与测试数据分布相近,未标注数据可以帮助模型提高分类精度。当数据分布相差较大时,会产生负迁移的现象。同时,测试集中抽取的数据信息不具有代表性,分类性能下降。实验表明,未标注的数据分布对应用于跨语言任务的迁移学习方法影响非常大。文献[6]中提出的方法通过情感词的上下文情感的模糊信息表示,判断模糊限制句来表达特定语境下的语义和情感信息的融合。该方法以情感词为中心,一方面可增强句子的情感表示,另一方面也受限于情感词的分布。没有充分利用语料本身的语义表达,所以分类效果有限。

相比之下,本文没有用到情感词典,仅用测评任务中给出的文本语料就取得了最高的准确率。且对三个产品类的评论分类精度都比较高,未曾出现不平衡的现象。这是因为本模型可以从抽象层面上提取了不同语种的语义特征,特征之间的隐含联系对分类的效果提高非常有帮助。且本模型的扩展度高,不仅适用于跨语言任务,还适用于多语种特征任务。从实验结果可以得知,基于联想记忆模型的CNN分类模型对跨语言情感分类非常高效。

4 结论

本文一方面分析了 dropout 大小对分类器性能的影响;另一方面对跨语言文本信息的提取方法与现有的方法进行了实验结果对比,证明了 dropout 的选取可以有效提高普通的分类器对短文本语义的识别度。通过自联想记忆关系对跨语言短文本语义进行融合,减少了在语言交换过程中语义的丢失,加强了短文本语义的完整表示程度。与其他学者提出的方法相比,本文设计的跨语言情感分类模型获得了非常好的效果,验证了本文所提出方法的有效性。

此次训练只针对了电商网站上的商品评论,下一步研究计划将该应用扩展到其他应用环境中,进

一步验证算法的有效性。

参考文献

- [1] Gui L, et al. Amixed model for cross-lingual opinion analysis[C]//Proceedings of Natural Language Processing and Chinese Computing, 2013: 93-104.
- [2] Gui L, et al. Cross-lingual opinion analysis via negative transfer detection[C]//Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, 2014: 860-865.
- [3] Chen Q, et al. Learning to adapt credible knowledge in cross-lingual sentiment analysis[C]//Proceedings of the Meeting of the Association for Computational Linguistics and the International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, 2015: 419-429.
- [4] Wang Y, Huang M, Zhao L. Attention-based lstm for aspect-level sentiment classification[C]//Proceedings of the 2016 conference on empirical methods in natural language processing, 2016: 606-615.
- [5] Li Z, et al. End-to-end adversarial memory network for cross-domain sentiment classification[C]//Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI 2017), 2017.
- [6] 陈龙. 基于深度学习的跨语言信息抽取研究[D]. 大连理工大学硕士学位论文. 2016.
- [7] Gliozzo A M, Strapparava C. Exploiting comparable corpora and bilingual dictionaries for cross-language text categorization[C]//Proceedings of the International Conference on Computational Linguistics and Meeting of the Association for Computational Linguistics (ACL), 2006: 553-560.
- [8] Hanneman G, Lavie A. Automatic category label coarsening for syntax-based machine translation[C]//Proceedings of the Workshop on Syntax, Semantics and Structure in Statistical Translation(ACL), 2011: 98-106.
- [9] Faruqi M, Dyer C. Improving vector space word representations using multilingual correlation[C]//Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics, 2014: 462-471.
- [10] Guo J, et al. Cross-lingual dependency parsing based on distributed representations[C]//Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, 2015: 1234-1244.
- [11] Harris Z S. Mathematical structures of language

- [M]. Interscience Tracts in Pure and Applied Mathematics. New York: Wiley, 1968: 111-112.
- [12] Jiao L, Rongyi C, Yahui Z. Cross-lingual similar documents retrieval based on co-occurrence projection [C]//Proceedings of the 6th International Conference on Computer Science and Network Technology, 2017:11-15.
- [13] Mikolov T, et al. Efficient estimation of word representations in vector space [J/OL]. [7 Sep 2013], arXiv:1301.3781v3.
- [14] Mikolov T, V L Q, Sutskever I. Exploiting similarities among languages for machine translation [J/OL]. [17 Sep 2013]. arXiv: 1309.4168v1.
- [15] Kim Y. Convolutional neural networks for sentence classification [C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014: 1746-1751.
- [16] Ioffe S, Szegedy C. Batch Normalization: Accelerating deep network training by reducing internal covariate shift [C]//Proceedings of the 32nd International Conference on Machine Learning, 2015:448-456.



刘娇(1994—), 硕士, 主要研究领域为多语种文本分析。

E-mail: mykongkongliu@foxmail.com



赵亚慧(1974—), 通信作者, 副教授, 主要研究领域为文本分析。

E-mail: yhzhaoy@ybu.edu.cn



崔荣一(1962—), 博士, 教授, 主要研究领域为智能计算, 模式识别, 机器学习, 自然语言处理。

E-mail: cuirongyi@ybu.edu.cn