

一种基于联合深度神经网络的食物安全信息情感分类模型

刘金硕 张 智

(武汉大学计算机学院 武汉 430072)

摘要 针对因中文食品安全文本特征表达困难,而造成语义信息缺失进而导致分类器准确率低下的问题,提出一种基于深度神经网络的跨文本粒度情感分类模型。以食品安全新闻报道为目标语料,采用无监督的浅层神经网络初始化文本的词语级词向量。引入递归神经网络,将预训练好的词向量作为下层递归神经网络(Recursive Neural Network)的输入层,计算得到具备词语间语义关联性的句子特征向量及句子级的情感倾向输出,同时动态反馈调节词向量特征,使其更加接近食品安全特定领域内真实的语义表达。然后,将递归神经网络输出的句子向量以时序逻辑作为上层循环神经网络(Recurrent Neural Network)的输入,进一步捕获句子结构的上下文语义关联信息,实现篇章级的情感倾向性分析任务。实验结果表明,联合深度模型在食品安全新闻报道的情感分类任务中具有良好的效果,其分类准确率和F1值分别达到了86.7%和85.9%,较基于词袋思想的SVM模型有显著的提升。

关键词 联合神经网络模型,多粒度文本特征,词向量,食品安全,情感倾向性分析

中图法分类号 TP391 文献标识码 A DOI 10.11896/j.issn.1002-137X.2016.12.051

Sentiment Analysis on Food Safety News Using Joint Deep Neural Network Model

LIU Jin-shuo ZHANG Zhi

(Computer School, Wuhan University, Wuhan 430072, China)

Abstract Facing the difficulties in feature expression of Chinese food safety text information, and the loss of semantic information with low classification accuracy, a sentimental text classification model based on joint deep neural network was presented. The proposed model utilizes the corpora of food safety document captured from the internet, and word vector from word embedding method as the input for the neural network to get the pre-trained word vector. The pre-trained word vector is further trained dynamically to get the word features and the sentimental classification of the sentence result, which better express the phrase-level sentimental relations for each sentence and the real semantic meaning in the food safety domain. Then the word feature of the sentence is inputted to the recurrent neural network (RNN) to catch the semantic information of the sentence structure further, realizing the sentimental classification of the text. The experiments show that our joint deep neural network model achieves better results in sentiment analysis on food safety information, compared with the bag-of-words based SVM model. The classification accuracy and F1 value reach 86.7% and 85.9% respectively.

Keywords Joint deep neural networks model, Multi-dimensional textual features, Word-embedding, Food safety, Sentiment analysis

1 引言

随着互联网以用户为中心的开放式架构的蓬勃发展,文本情感分类技术在用户评论分析与决策、舆情监控、电子政务、推荐系统以及趋势预测等多个应用领域均做出了巨大贡献^[1]。文本情感分类技术日臻成熟,特征表示成为研究的技术瓶颈,人工或者无监督的特征选择的好坏决定了分类模型的优劣。根据词频信息、情感约束规则和否定词前缀等特征影响因子进行选择优化,并在此基础上构建情感词典及进行句法结构分析,使有标记的数据的作用最大化,有效地提高了传统线性分类模型的准确率^[2]。已有研究中基于词袋思想的人工特征选择难以在分类过程中发掘潜在的情感倾向相关性语义特征,如瘦肉精、三聚氰胺、地沟油等食品安全领域常用

的负向词语,在特征表示的时候需要大量人工标记的数据做支持,并且在特定领域内根据已标注的文本进行未标注文本学习的难度极大。

本文提出一种联合深度神经网络模型,该模型以浅层网络预训练的词向量为输入,捕获词语和句子两级文本粒度之间的相互关联性及其隐含语义特征,实现篇章级的文本情感倾向判别。首次提出采用联合神经网络的方法处理食品安全新闻信息的情感分类问题,并获得了正负二元标注的食品安全信息语料库。通过实验发现,两级全连接的神经网络在分类过程中逐层反馈调节各节点单元向量,使单元向量更加接近其真实语义环境下的表达,明显提高了分类的准确率。

本文第2节介绍传统情感分类方法和基于深度学习的情感倾向性分析的相关工作;第3节介绍本文提出的基于联合

到稿日期:2015-11-26 返修日期:2016-02-25 本文受国家自然科学基金(61303214)资助。

刘金硕(1973-),女,博士,副教授,主要研究方向为模式识别、自然语言处理、数字图像处理,E-mail:liujinshuo@whu.edu.cn;张智(1990-),男,硕士,主要研究方向为自然语言处理,E-mail:zhangzhi@whu.edu.cn(通信作者)。

深度神经网络的食品安全信息分类方法;第4节为实验及实验结果分析;最后总结全文。

2 相关工作

2.1 传统情感分类

2002年,康奈尔大学的 Bo Pang 等人^[3]首次提出了文本情感分类的概念,并采用贝叶斯(Naive Bayes)、最大熵和支持向量机(SVM)等模型^[4]进行文本情感倾向预测实验,为自然语言文本倾向性分析的研究拉开了序幕。半监督的情感分类技术以其兼顾人工标注的成本和模型自主学习效果的优势,成为近年来传统分类器进行模型优化的研究热点。Yang 和 Cardie 等^[5]将情感分类任务分解为弱监督的情感片段识别和有监督的属性标注两个子过程,分别建立概率模型并进行联合训练,通过获取不同粒度的文本情感表达式之间的相互影响关系和情感关联性,完成句子的情感标签得分计算。线性分类模型在语义捕获方面的劣势使其在没有巨大人工投入的前提下,在分类效果上始终存在难以突破的瓶颈。

中文微博数据的情感分类研究一直是该领域内的研究热点。张志琳等^[6]提出了基于词汇化主题特征、情感词内容特征和概率化的情感词倾向性等多样化特征的中文微博情感分类方法,显著改善了微博情感分析的性能。梁军等^[7]探讨了基于深度学习的方法来进行中文微博的情感倾向性分析的可能性,采用神经网络来发现任务相关的特征,避免依赖于手工特征设计,节省了大量人工标注的工作量。微博短文本的情感分类研究给篇章级长文本分类任务带来了启发,篇章级文本具备更加丰富的语义信息,短文本的特征选择方法在面对多级粒度结构的长文本时,难以捕获全局的语义信息,使分类效果达不到预期的效果。

2.2 基于深度学习的情感倾向性分析

深度学习的动机在于模拟人脑进行分析学习,将信息通过简单非线性的多层神经网络叠加的模型转变成更高层次、更加抽象的表达,它在图像处理和语音识别方向的成功应用给自然语言情感倾向性分析的研究带来了新的启发。Maas 等^[8]提出了一种基于词向量的情感分类方法,该方法采用无监督的模型计算词语之间的语义相似度,同时以有监督的方法获取文本中存在的语义信息,模型在标注的公共语料库的分类实验中取得了非常不错的效果。词语的极性计算在一定程度上影响着分类效果,但分类模型在分类任务中起着更为关键的作用。Socher 等^[9]基于递归的自动编码器给出了一个新型的机器学习框架,该框架学习词语和短语块等不同粒度情感表达式的向量表示,并将学习结果作为情感分类器的特征输入,判断句子级的文本倾向性。随后, Mikolov 等^[10]对神经网络模型进行了优化,通过神经网络隐含层的反复利用,不断与新加入的时序信息联合计算出下一个隐含层,这样的网络模型也被称为循环神经网络(Recurrent Neural Network)。隐含层包含了所有的上下文语义信息,同时有效地减少了模型的参数数量,极大地降低了模型计算的复杂度。

早期的神经网络模型采用稀疏存储的词向量^[11]作为输入,随着网络层次的增加,高维的特征使模型的参数学习成为了灾难,并且相似词语之间不具备关联性,导致在模型训练时出现语义缺失问题。Mikolov^[12,13]在其随后的研究工作中使用浅层神经网络模型将词语转换成词向量,并推出了词向量转换工具 Word2Vec。使用深度学习的方法进行篇章级的情感分析,词语到篇章之间存在文本跨度,因而涉及到多级

粒度的语义信息的计算,单一的网络模型并不能解决全局语义缺失的问题。本文提出的联合深度模型分别以递归和循环的计算模式获取多级粒度文本之间的情感关联信息和语义特征,提高了深层语义的情感分析效果。

3 联合深度神经网络分类模型

本文提出的基于联合神经网络的跨文本粒度情感倾向性分析模型如图1所示。该模型由底层的递归神经网络(RNN)和上层的循环神经网络(RcNN)联合构成。

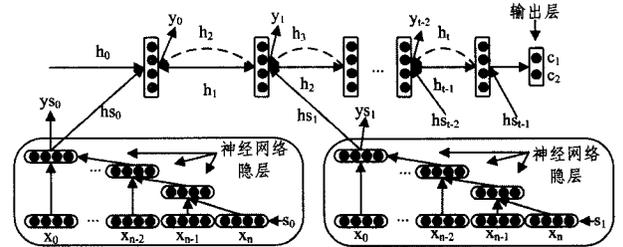


图1 基于联合神经网络的跨文本粒度情感倾向性分析模型

(1)文本向量化。分别通过中文分词技术和基于标点的方法将篇章语料划分为词语 x 和句子 s 两个粒度级别,并用一个浅层的神经网络将词语表示为一个低维的实数向量 x_i ,词语向量与整个词表向量矩阵形成映射关系,即 $x_i \in R^n, R^n$ 为对应向量矩阵。

(2)基于递归神经网络的句子级情感分析。把预训练好的词向量 x_i 作为递归神经网络的输入,经过输入层网络节点计算得到隐含层向量元素,然后使用词向量与隐含层向量生成新的隐含单元,递归使用隐含层最终获得句子向量 h_s 。在层与层节点计算的过程中加入矩阵因子,使网络中所有的节点都具有相同的维度,避免维数灾难问题。最后,在网络顶层加入分类器节点,预测句子的情感倾向 y_s 。

(3)基于循环神经网络的篇章级情感分析。句子向量 h_s 以所在篇章的位置结构先后被递归神经网络计算出来,与循环神经网络的前一隐含节点 h_{i-1} 生成当前隐含节点 h_i ,隐含节点被反复循环利用,最新的隐含节点包含了句子关联语义信息。这样,词语、中间短语、句子及段落等多级粒度的语义特征表示均接近其真实表达。同样,在循环神经网络的顶层节点加入分类器单元计算情感得分 y_i ,并判断其二元倾向 c_i 。

3.1 基于递归神经网络的句子级情感分析

词向量以固定的维度被训练得出,不同句子所包含的词向量数目不一致,采用拼接或者平均值的处理会造成高级别的向量表示维度不一,从而损失部分文本语义信息。本文采用递归神经网络模型处理食品安全新闻语料,在实现句子级情感倾向判别的同时,输出同等维度的句子单元向量。同时,在每个递归神经网络的顶层加入 sigmoid 回归函数,计算句子的文本情感倾向,如图2所示。

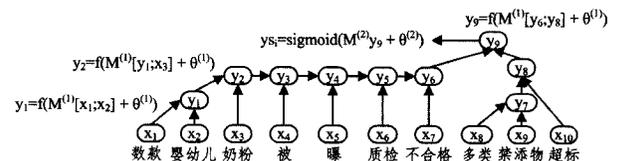


图2 句子级情感分析网络模型

$x = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}$ 构成了图2中句子的词向量集,成为神经网络的输入层,递归模型中的基本单元为三元结构 $p \rightarrow n_1 n_2$, p 为生成单元节点,子节点为输入词向量 x_i 或者中间隐含向量 y_i ,如图2中的树结构($y_1 \rightarrow x_1 x_2$),

$(y_2 \rightarrow y_1, x_3), (y_9 \rightarrow y_6, x_8)$, 加入向量矩阵 $M^{(1)}$ 之后, 模型中所有节点向量的维度被递归统一。

$$h_{s_i} = f(M^{(1)}[n_1; n_2] + \theta^{(1)}) \quad (1)$$

其中, $\theta^{(1)}$ 为偏置, 矩阵参数 $M^{(1)} \in R^{d \times 2d}$, d 表示词向量及中间隐含单元的维度, 在完成高粒度级别向量降维的同时, 中间隐含单元并不损失任何语义信息。

叶子节点层的词向量被逐层反馈调节, 生成同等维度且具有词语级全局特征信息的句子单元向量 h_{s_i} , 分类器根据 h_{s_i} 计算句子的二元情感倾向得分 y_{s_i} , 如式(2)所示:

$$y_{s_i} = \text{sigmoid}(M^{(2)} h_{s_i} + \theta^{(2)}) \quad (2)$$

3.2 基于循环神经网络的篇章级情感分析

深度模型在篇章级情感分类任务中由于文本结构的问题, 以词向量作为输入的普通神经网络模型难以捕获长距离语义信息。循环神经网络隐含单元相互连接, 信息循环传递, 使其具备获取任意长度上下文语义信息的能力。篇章级的文本计算量巨大, 为避免模型参数学习困难, 采用单隐层循环神经网络框架, 其隐层和输出层的计算规则如下:

$$h_j(t) = f(\text{RcNN}_j(t)) \quad (3)$$

$$\text{RcNN}_j(t) = \sum_i^m h_{s_i}(t) v_{ji} + \sum_k^n h_k(t-1) u_{jk} + \theta^{(3)} \quad (4)$$

$$y_j(t) = g(\sum_j^m h_j(t) w_{kj} + \theta^{(4)}) \quad (5)$$

其中, h_j 是神经网络的第 j 个隐含单元, RcNN 表示上层循环神经网络计算单元, $h_j(t)$ 为底层递归模型计算生成的句子单元向量, 每个句子单元向量被看作循环神经网络的 t 时刻的输入; u 和 v 为输入层到隐层的向量矩阵, w 为隐层到输出层的向量矩阵, $\theta^{(3)}$ 和 $\theta^{(4)}$ 为偏置。 $f(x)$ 和 $g(x)$ 分别为激活函数和分类函数, 由于旨在实现篇章二元分类, 因此:

$$f(x) = g(x) = \text{sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (6)$$

模型当前时间步长的隐含单元 $h_j(t)$ 由当前输入 $h_{s_i}(t)$ 和前一时间步长隐含单元 $h_j(t-1)$ 经激活函数生成。当 $t=1$ 时, $h_j(0)$ 被随机初始化为全部特征值为 0.01 的向量, 最小化了人为选择对模型的影响。

3.3 联合模型训练

令 $\delta = (M^{(1)}, \theta^{(1)}, M^{(2)}, \theta^{(2)}, v_{ji}, u_{jk}, w_{kj}, \theta^{(3)}, \theta^{(4)})$ 为联合神经网络模型的参数集合, 则给定文本的情感标签 t 和参数集 δ 的概率分布为:

$$y(t; \delta) = \text{sigmoid}(M^{(3)} y(t) + \theta^{(5)}) \quad (7)$$

其中, $y(t)$ 为经过联合网络模型计算而具备全局语义信息的篇章特征向量, $M^{(3)}$ 为篇章级标签向量矩阵, 对应的似然函数为:

$$d(t; \delta) \log(y(t; \delta)) + (1 - d(t; \delta)) (1 - \log(y(t; \delta))) \quad (8)$$

将似然函数记为 $L(\delta)$, 使用交叉熵进行误差计算, 得到的代价函数为:

$$J(\delta) = -\frac{1}{m} \sum_{(x,s)} L(\delta) + \frac{\lambda}{2} \alpha_j^2 \quad (9)$$

其中, $d(t; \delta)$ 是期望输出, $y(t; \delta)$ 为模型计算输出; m 是上层神经网络输出节点总数, 表示所有节点对最终分类效果都有贡献; $(x; s)$ 为已标注的(句子, 情感标签)序列对, $\frac{\lambda}{2} \alpha_j^2$ 为惩罚因子, 其作用是增强模型泛化能力。联合模型使用随机梯度下降^[14]最小化似然函数:

$$\frac{\partial J(\delta)}{\partial \delta} = -\frac{1}{m} \sum_{(x,s)} \frac{\partial L(\delta)}{\partial \delta} + \lambda \alpha \quad (10)$$

在模型参数学习的过程中由于参数过多, 为了简化梯度

计算的复杂度, 矩阵参数均使用局部最优解, 并不随梯度计算一直调整。同时, 将循环神经网络设计为浅层结构, 以降低梯度计算过程中的误差传播。

4 实验结果和分析

4.1 数据集信息

本文实验采用的数据集是利用分布式网络爬虫分别从新华网食品新闻舆情在线、人民网食品频道、中国食品安全论坛官方网站和中国食品安全网 4 个门户网站获取的 2013 年 3 月 20 日到 2015 年 10 月 1 日的全部新闻报道, 根据标题进行去重处理之后, 共有 159572 条文本。在信息获取的过程中, 按月对当前月的所有文本进行人工二元情感标注。最后, 实验语料中负向文本的数量为 36469 条, 由于本文旨在实现食品安全新闻信息的二元分类任务, 其他的文本均作为非负文本处理。以时间刻度将文本集划分为训练集、验证集以及测试集, 详细的统计信息如表 1 所列。

表 1 数据分布统计信息

	训练集	验证集	测试集
文档	126412	18483	14677
句子	2654676	406641	348235
时间跨度	2013/3/20~ 2014/12/20	2014/12/21~ 2015/5/22	2015/5/23~ 2015/10/1

4.2 对比实验

本文设计了多个实验来验证联合深度模型的有效性。在传统 SVM 线性分类模型中, 输入的是 Bi-gram 词袋向量, 特征值使用 TF-IDF 计算; 而对于所有的深度模型, 词向量的维度为 200, 隐含神经元个数为 300, 使用 Word2Vec 预训练词向量的窗口大小为 5, 即每个词具有左右各 5 个词的上下文信息。

BOW-SVM: 词袋向量为初始输入, 训练过程中人工参与特征选取, 使用 SVM 分类器进行分类;

RV-RNN: 词向量随机初始化作为递归神经网络的输入, 神经网络反馈调节词向量, 采用二元分类器进行句子级的情感分类;

Static-WE-RNN: 使用 Word2Vec 训练词语词向量, 将其结果作为递归神经网络的输入, 词向量在实验过程中保持不变, 采用二元分类器进行句子级的情感分类;

Dynamic-WE-RNN: 使用 Word2Vec 训练词语词向量, 将其结果作为递归神经网络的输入, 词向量受神经网络反馈调节, 采用二元分类器进行句子级的情感分类;

SE-RcNN: 使用预训练好的词向量做平均值处理得到句子向量, 将其作为循环神经网络的输入, 进行篇章级的情感分类任务;

RNN-RcNN: 本文所提方法, 即采用 Word2Vec 预训练词向量, 将其结果作为底层递归神经网络的输入, 生成中间隐含句子单元向量, 输入到上层循环神经网络, 在整个网络模型中, 各节点单元向量在分类模型的训练过程中受神经网络反馈进行微调, 最后完成篇章级文本分类的任务。

4.3 结果和分析

在分类实验的过程中根据不同模型的特点, 分别以标题、标题加正文两种粒度的文本语料进行实验, 实验结果如表 2 所列。

从表 2 的实验结果中, 可以发现:

(1) 基于词袋向量的 SVM 分类模型(BOW-SVM)需要进

行人工特征选取,大量潜在的上下文语义信息并不能被表示出来,导致分类的结果并不理想。为了解决只以标题为语料带来的数据量不足的问题,本文在标题的基础上加入正文,分类结果并没有大幅改善,原因有可能是食品安全新闻正文中存在大量的噪声因子,干扰了分类器分类的效果。

(2)词向量在句子级的分类任务中有着十分重要的作用,对于相同的情感分类模型(RNN),采用 Word2Vec 预训练词向量作为输入的模型(Static-WE-RNN)相较于随机初始化的词向量的模型(RV-RNN)在分类准确率上有明显提升(85.2% vs 83.5%)。Word2Vec 基于一种浅层的神经网络生成词向量,能够捕获词与词之间的关联性及相关语义信息,使词向量更接近词语本身的真实表达。

(3)在递归神经网络的训练过程中,输入的词向量受网络反馈调节(Static-WE-RNN vs Dynamic-WE-RNN)准确率有进一步的提升(0.6%)。已具备语义信息的词向量在模型参数学习过程中,增加了全局的语义特征信息,更加适应于特定领域内的语料,进一步提高了分类准确率。

(4)新闻语料的标题概括性比较大,在一定程度上可以弥补上下文语义信息的不足,但是以标题为语料的模型始终突破不了语义信息不足的瓶颈。采用词向量平均值的方法构造的句子单元向量作为篇章分类神经网络模型的输入(SE-RcNN),虽然在实验中句子向量被反馈调节,但是损失了句子内部词之间的语义信息,分类效果(82.4% vs 86.7%)远不如以全连接神经网络捕获包括词级别与句子级别全局语义特征信息的分类模型(RNN-RcNN)。

表2 不同分类模型的结果对比

模型	语料	准确率(%)	召回率(%)	F1值(%)
BOW-SVM	标题	78.7	76.3	77.5
BOW-SVM	标题+正文	79.1	77.4	78.2
RV-RNN	标题	82.5	81.0	81.7
Static-WE-RNN	标题	84.0	82.6	83.2
Dynamic-WE-RNN	标题	84.6	83.1	83.8
SE-RcNN	标题+正文	82.4	81.5	81.9
RNN-RcNN	标题+正文	86.7	85.2	85.9

同时,如图3所示,本文还在 RNN-RcNN 模型中针对不同维度的词向量对分类结果的影响进行了实验研究。实验结果表明,在海量数据支持的前提下,词向量的维度越高,其具备越高的精度和描述不同词之间关联性的能力,最终的结果效果也越好。在已有语料的基础上进行实验,发现当词向量的维度超过 200 时,分类的准确率仅有小幅的提升,但是模型训练的复杂度会急剧上升,因此根据实验结果,最终的模型采用的词向量维度为 200。

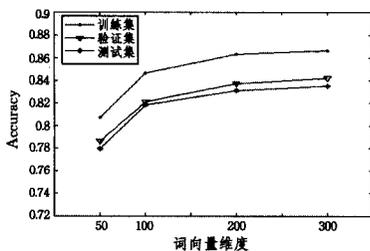


图3 词向量维度对分类准确率的影响

结束语 本文为避免依赖于特定任务的人工特征设计,采用神经网络来发现与文本情感分类相关的特征,结合多级文本粒度的情感相互关联性,提出了一种基于联合深度神经网络的情感倾向性判别方法,实现了食品安全领域新闻类篇章级的情感分类任务。将词向量作为递归神经网络和循环神

经网络全连接模型的初始输入,以递归计算的形式得到高级别粒度的句子向量表示,并在上层网络中捕获全局多级粒度的语义信息。实验表明,本文方法可有效解决特定领域内长文本情感分类任务中因特征表达困难造成的语义缺失等问题,提高了情感分类的准确率。接下来的研究重点是采用基于联合深度模型的方法,解决特定领域内多元细粒度的情感分类问题,同时进行模型优化。

参考文献

- [1] Zhao Yan-yan, Qin Bing, Liu Ting. Sentiment Analysis [J]. Journal of Software, 2010, 21(8): 1834-1848 (in Chinese)
赵妍妍, 秦兵, 刘挺. 文本情感分析 [J]. 软件学报, 2010, 21(8): 1834-1848
- [2] Pang B, Lee L. Opinion mining and sentiment analysis [J]. Foundations and Trends in Information Retrieval, 2008, 2(1/2): 1-135
- [3] Pang B, Lee L. Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales [C] // Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics, 2005: 115-124
- [4] Pang B, et al. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts [C] // Proceeding of ACL. 2004: 271-278
- [5] Yang B S, Cardie C. Joint Modeling of Opinion Expression Extraction and Attribute Classification [J]. Transactions of the Association for Computational Linguistics, 2014, 2(1): 505-516
- [6] Zhang Zhi-lin, Zong Cheng-qing. Sentiment Analysis of Chinese Micro Blog Based on Rich-features [J]. Journal of Chinese Information Processing, 2015, 29(4): 134-142 (in Chinese)
张志琳, 宗成庆. 基于多样化特征的中文微博情感分类方法研究 [J]. 中文信息学报, 2015, 29(4): 134-142
- [7] Liang jun, Chai Yu-mei, Yuan Hui-bin, et al. Deep Learning for Chinese Micro-blog Sentiment Analysis [J]. Journal of Chinese Information Processing, 2014, 28(5): 155-160 (in Chinese)
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): 155-160
- [8] Maas A L, Daly R E, Pham P T, et al. Learning word vectors for sentiment analysis [C] // Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics; Human Language Technologies, 2011, 1: 142-150
- [9] Socher R, Perelygin A, Wu J Y, et al. Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank [C] // Proceedings of the Empirical Methods in Natural Language Processing, 2013: 1631-1642
- [10] Mikolov T, Zweig G. Context Dependent Recurrent Neural Network Language Model [J]. IEEE Spoken Language Technology Workshop, 2012, 8537(11): 234-239
- [11] Hinton G. E, et al. Learning Distributed Representations of Concepts [C] // Proceedings of the 8th Annual Conference of the Cognitive Science Society, 1986
- [12] Milkolov T, Kombrink S, et al. Extensions of Recurrent Neural Network Language Model [J]. IEEE International Conference on Acoustics, Speech & Signal Processing, 2011, 125(3): 5528-5531
- [13] Milkolov T, Chen K, Corrandi G, et al. Efficient estimation of word representations in vector space [J]. arXiv:1301.3781
- [14] Boden M. A Guide to Recurrent Neural Networks and Back-propagation [J]. Dallas Project Sics Technical Report T Sics, 2001, 36(6): 112-119