计算机系统应用 ISSN 1003-3254, CODEN CSAOBN
Computer Systems & Applications,2018,27(8):153−158 [doi: 10.15888/j.cnki.csa.006480]

E-mail: csa@iscas.ac.cn
http://www.c-s-a.org.cn
Tel: +86-10-62661041

基于改进词向量的石油文档语义关系识别①

宫法明, 朱朋海

(中国石油大学 (华东) 计算机与通信工程学院, 青岛 266580)
通讯作者: 朱朋海, E-mail: z15070507@s.upc.edu.cn

摘　要: 语义关系识别是对文档进行处理识别出包含的语义关系的过程, 是构建本体重要组成部分之一. 在石油领域本体的构建过程中, 由于石油领域的文档具有组合词多的特点, 语义关系识别更加困难. 目前使用的语义识别算法主要是基于关联规则的识别算法, 但此类算法没有领域针对性. 通过分析石油文档的特点, 提出一种基于改进词向量的石油文档语义关系识别算法, 以连续词袋 (Continuous Bag-Of-Words, CBOW) 模型为基础, 对石油专业术语进行扩展训练, 引入负采样和二次采样技术提高训练准确率和效率, 利用向量特征训练支持向量机 (Support Vector Mechine, SVM) 分类器进行语义关系识别. 实验结果表明, 该方法训练的词向量能够准确识别石油领域的语义关系, 在石油领域具有明显的优势.

关键词: 词向量; 语义关系识别; SVM

引用格式: 宫法明,朱朋海.基于改进词向量的石油文档语义关系识别.计算机系统应用,2018,27(8):153–158. http://www.c-s-a.org.cn/1003-3254/6480.html

Semantic Relationship Recognition of Oil Documents Based on Improved Word Vector

GONG Fa-Ming, ZHU Peng-Hai

(College of Computer &Communication Engineering, China University of Petroleum, Qingdao 266580, China)

Abstract: Semantic relationship recognition is the process of document processing and is used to identify the semantic relations contained in the process, which is an important part of the construction of ontology. In the process of constructing petroleum field ontology, the semantic relationship identification is more difficult because the documents in the petroleum field have their unique characteristics. The current semantic recognition algorithm is mainly based on association rules' recognition algorithm, but there is no field-specific orientation. By analyzing the characteristics of petroleum documents, this study proposes a semantic relationship recognition algorithm for petroleum documents based on improved word vector. Based on the Continuous Bag-Of-Words (CBOW) model, this study carries out expanded model training on petroleum terminologies and introduces negative sampling and subsampling techniques to improve the training accuracy and efficiency. Feature vectors are used in training the Support Vector Mechine (SVM) classifier for semantic relationship recognition. The experimental results show that the word vectors trained by this method can accurately identify the semantic relations contained in documents in the petroleum field and have obvious advantages.

Key words: word vector; semantic relationship recognition; Support Vector Mechine (SVM)

1　引言

语义是指信息包含的概念和意义. 语义不仅表述事物本质, 还表述事物之间的因果、上下位、施事等各种逻辑关系. 因此, 语义是对事物的描述和逻辑表示.

① 基金项目: 科技部创新方法工作专项 (2015IM01030)
Foundation item: Special Project for Innovation Work of the Ministry of Science and Technology of China (2015IM01030)
收稿时间: 2017-12-10; 修改时间: 2018-01-04; 采用时间: 2018-01-16; csa 在线出版时间: 2018-07-28

语义分析就是对信息所包含的语义的识别, 并建立一种计算模型, 使其能够像人那样理解自然语言. 语义分析是自然语言理解的根本问题, 它在自然语言处理、信息检索、信息过滤、信息分类、语义挖掘等领域有着广泛的应用. 在互联网时代, 面对海量的信息资源, 要想准确地进行信息抽取, 检索所需信息、挖掘潜在的信息价值、提供智能的知识服务, 都离不开面向机器理解的语义分析. 尤其在大数据环境下, 语义分析的地位越来越凸显出来[1].

关于语义分析研究方法主要有基于专业词典的方法[2–4]、基于词汇-句法模式的方法[5-6]、基于 Harris 假设的方法[7–9]、基于关联规则的方法[10,11]、基于模式匹配的方法和几种方法的混合方法. 但这些方法存在很大的局限性, 无法涵盖所有的内容, 不能针对石油领域组合词多的特点进行语义识别, 没有领域适应性.

为了克服以上算法的缺点, 本文提出了一种基于改进词向量的语义关系识别模型, 该模型克服了传统词向量训练模型固定大小的缺点, 考虑到句子的完整性对词向量训练的重要作用. 具体地, 以 CBOW 模型为基础, 在输入层前增加一个预处理层, 对句子进行预处理, 转换成统一输入格式, 方便进行训练. 为了适应石油领域组合词以及具有父子关系的词多的特点, 对训练语料进行扩展, 增加正样本训练比重, 使训练结果更加准确. 通过负采样和二次采样提高训练效率和训练精度. 将训练好的词向量进行代数运算提取特征进行 SVM 分类识别出语义关系. 对样本的准确率、召回率和 F 值进行计算分析, 与传统的 CBOW 模型相比, 我们的语义关系识别效果更好. 本文贡献如下所示:

(1) 提出了一种适用石油领域改进的词向量训练方法;

(2) 本文提出的词向量能够准确识别石油文档中的类义关系;

(3) 为石油文档语义识别提供了一个思路.

本文章安排具体如下: 第 2 节介绍相关理论与方法; 第 3 节介绍改进的词向量表示方法; 第 4 节是实验结果及分析; 最后对本文工作进行了总结, 并指出为未来的工作方向.

2 相关工作

基于词向量的文档语义识别, 通过对语料库中的文本进行分词处理, 统计各个词语出现的频率, 采用 one-hot 表示方法初始化词向量, 构建词语矩阵, 利用矩阵变换训练调整词向量. 该算法有简单方便、易于训练、效果好等优点, 但也有领域适应性差、计算复杂、效率低等缺点. 周慧霞[12]提出了一种基于词向量的中文词汇蕴含关系识别方法, 利用词向量技术对维基百科上的的语料进行训练, 设计词向量分类特征进行词汇蕴含关系的识别. 由于石油领域组合词多, 该方法领域适用性差, 并不适用石油领域, 并且也不能识别上下义关系. 蒋振超[13]等提出了一种新的基于神经网络的词向量训练模型, 借助大规模文本数据, 利用依存关系和上下文关系来训练词向量. 但此方法需要依据大规模文本数据, 由于石油行业的保密性较强, 无法提供充分的文本数据进行训练. 杜漫[14]等提出了一种面向情绪分类的融合词内部信息和情绪标签的词向量学习方法. 在 CBOW 模型基础上, 引入词内部成分和情绪标签信息, 以适应微博情绪表达的不规范, 但这种方法只针对微博信息进行处理, 领域性太强, 不适合石油领域.

尽管他们能够取得比较好的语义关系识别结果, 但都没有领域针对性, 对领域专业词之间的语义关系识别效果并不太好. 在本文中, 我们提出了一种基于改进词向量的石油文档语义识别模型算法, 该算法能够很好的识别领域专业词语之间的语义关系, 特别是针对石油领域这样组合词多的领域.

3 改进的词向量训练模型

3.1 CBOW 模型简介

CBOW 模型是常用的词向量训练模型, CBOW 模型的训练输入是一个特征词的上下文相关词对应的词向量, 而输出就是这特定词的词向量. CBOW 模型的结构如图 1, 该模型一方面根据 C&W[15]模型的经验, 使用一段文本的中间词作为目标词; 另一方面, 又以 NNLM[16,17]作为蓝本, 并在其基础上做了两个简化.

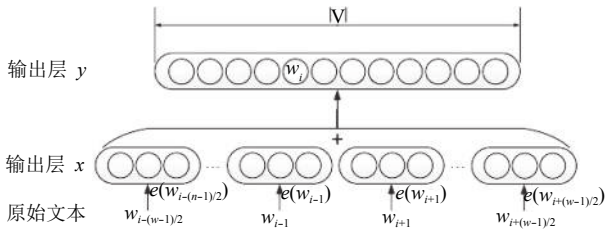

图 1 CBOW 模型结构图

(1) CBOW 没有隐藏层, 去掉隐藏层之后, 模型从神经网络结构直接转化为 log 线性结构, 与 Logistic 回

归一致. log 线性结构比三层神经网络结构少了一个矩阵运算, 大幅度地提升了模型的训练速度.

(2) CBOW 去除了上下文各词的词序信息, 使用上下文各词词向量的平均值, 代替神经网络语言模型使用的上文各词词向量的拼接. 形式化地, CBOW 模型对于一段训练样本$w_{i-(n-1)},\cdots,w_i$, 输入为:

$$x=\frac{1}{n-1}\sum_{w_j\in c}e(w_j) \tag{1}$$

由于没有隐藏层, CBOW 模型的输入层直接就是上下文的表示. CBOW 模型根据上下文的表示, 直接对目标词进行预测:

$$P(w|c)=\frac{\exp(e'(w)^{\mathrm{T}}x)}{\sum_{w'\in V}\exp(e'(w')^{\mathrm{T}}x)} \tag{2}$$

上述二式中, 目标词 w 和上下文 c 的定义与 C&W 模型中的公式一致. e 为词语 w 的词向量. 对于整个语料而言, 与神经网络语言模型类似, CBOW 的优化目标为最大化:

$$\sum_{(w,c)\in D}\log P(w|c) \tag{3}$$

3.2 改进的词向量训练方法

石油领域文档具有组合词和上下义关系词多的特点, 传统的 CBOW 训练方法固定上下文, 没有针对石油领域文档的特点进行训练, 不具有领域针对性, 对石油领域的词语训练结果较差. 本文提出的改进的词向量训练方法, 是在 CBOW 模型的基础上进行改进的. 例如“游梁式抽油机”、“抽油机”和“磕头机”等, “游梁式抽油机”、“抽油机”和“磕头机”具有相似的上下文语义结构, 所以训练“游梁式抽油机”、“抽油机”和“磕头机”采用的语料包括“抽机油”的上下文、“磕头机”的上下文和“游梁式抽油机”的上下文, 强化正样本. 其次, 词语包含在句子中, 每个词语受同一句中的词语影响较大, 受不同句中的词语影响较小. 对词语进行训练的时候, 选取包含词语所在句子为原则, 然后进行语料扩充, 而不是采用 CBOW 的固定上下文.

首先按照石油词典找出石油领域专业词语 w, 找出 w 的组合词$v_1,\cdots,v_m$. 选取中心词w_i所在的句子作为待处理语料, 获取句子长度 l 与中心词在句子中的位置 i, 计算中心词与句子两端的距离$l_1=i$, $l_2=l-i$, 训练窗口为n_w.

$$n_w=\begin{cases}2l_1+1 & l_1\geqslant l_2\\ 2l_2+1 & l_1<l_2\end{cases} \tag{4}$$

选取 w_i 上下文长度为n_w的词语$w_{i-(n_w-1)/2},\cdots,w_{i+(n_w-1)/2}$. 选取中心词$w_i$ 的组合词$v_1,\cdots,v_m$进行同样预处理, 获取$v_1,\cdots,v_m$ 上下文词语$w_{1i-(n_{v1}-1)/2},\cdots,w_{1i+(n_{v1}-1)/2},w_{2i-(n_{v2}-1)/2},\cdots,w_{2i+(n_{v2}-1)/2},\cdots,w_{mi-(n_{vm}-1)/2},\cdots,w_{mi+(n_{vm}-1)/2}$. 将$w_i$与$v_1,\cdots,v_m$的训练语料进行合并, 制作新的训练数据集$c=\{w_{i-(n_w-1)/2},\cdots,w_{i+(n_w-1)/2},w_{1i-(n_{v1}-1)/2},\cdots,w_{1i+(n_{v1}-1)/2},\cdots,w_{mi-(n_{vm}-1)/2},\cdots,w_{mi+(n_{vm}-1)/2}\}$, 公式 (1) 变成

$$x=\frac{1}{n_x+\sum_{i=1}^{m}n_{yi}-m-1}\sum_{w_j\in c}e(w_j) \tag{5}$$

3.3 负采样和二次采样技术

(1) 负采样技术

为了进一步提升最后一层的效率, 借鉴 C&W 模型采用的构造负样本的方法, 提出了负采样技术. 通过构造优化目标, 实现最大化正样本的似然, 同时最小化负样本的似然. 使用负采样, 对语料库进行排序, 通过随机选取一个较少数目的负样本来更新对应的权重, 并且仍然为正样本更新对应的权重, 如果选到自己就跳过重新选择.

(2) 二次采样技术

在大规模语料中, 高频词通常就是停用词 (如英语中的“the”、汉语中的“的”). 一方面, 这些高频词只能带来非常少量的语义信息, 比如几乎所有的词都会和“的”共同出现, 但是并不能说明这些词的语义都相似. 另一方面, 训练高频词本身占据了大量的时间, 但在迭代过程中, 这些高频词的词向量变化并不大. 为了解决这一问题, 提出了二次采样技术, 具体而言, 如果词 w 在语料中的出现频率 $f(w)$大于阈值 t, 则有$P(w)$的概率在训练时跳过这个词.

$$P(w)=1-\sqrt{\frac{t}{f(w)}} \tag{6}$$

3.4 语义关系特征

本文利用多个向量多个特征进行语义识别, 对于两个向量$u=<u_1,\cdots,u_n>$和$v=<v_1,\cdots,v_n>$, 如果两个词语具有相似关系, 则两个向量在同维度上的分量也具有相似关系, 他们的向量差就很小, 某一维度的分量和就很大. 各个向量特征定义如下:

向量差特征:

$$f_{diff}=u-v=<u_1-v_1,\cdots,u_n-v_n> \tag{7}$$

向量和特征:

$$f_{add}=u+v=<u_1+v_1,\cdots,u_n+v_n> \tag{8}$$

向量乘特征:

$$f_{mul} = u \cdot v = < u_1 \times v_1, \cdots, u_n \times v_n > \quad (9)$$

向量连接特征:

$$f_{cat} = < u, v > = < u_1, \cdots, u_n, v_1, \cdots, v_n > \quad (10)$$

3.5 改进的词向量语义识别算法实现

本算法是基于改进词向量模型实现的, 根据石油领域文档组合词多的特点, 实现适合石油领域词向量的训练模型, 利用词向量的加、减、乘和连接特征训练 SVM 分类器进行语义识别. 具体实现算法如下:

(1) 利用隐马尔科夫模型对文档 D 进行分词处理, 将文档分成单词的词语, 构建语料库 S;

(2) 对单词进行词频统计, 利用词频够将初始化词向量;

(3) 利用石油词典找出待训练的石油专业术语 w, 统计术语所在句子长度 l 和相关的组合词$v_1, \cdots, v_m$等信息;

(4) 利用 w 的上下文词语和$v_1, \cdots, v_m$的上下文词语, 结合负采样和二次采样对 w 进行词向量训练, 得到词语 w 的词向量;

(5) 对训练好的词向量进行向量差、和、乘和连接特征计算, 进行语义关系标记, 制作训练数据集;

(6) 将训练好的数据集导入 SVM 分类器进行语义关系训练;

算法流程图如图 2 所示.

4 实验结果及分析

本文基于中文维基百科语料库和 100 篇石油文档, 涵盖了油气地质勘探、油气田开发与开采等石油领域的 10 个子学科, 每个子学科 10 篇文档. 然后利用 3.5 节所述特征, 利用 libsvm 工具进行训练词汇语义关系 SVM 分类器, 最后在测试语料库上对识别效果进行综合评价.

4.1 语料库制作及试验模型参数设置

首先从维基百科的语料库中下载中文维基百科数据集并进行处理, 将 100 篇石油文档进行处理制作数据集, 然后用隐马尔科夫模型进行分词和词性标注, 制作成语料库. 然后利用 word2vec 对语料进行训练, 得到 50 维、100 维、200 维和 400 维词向量.

利用是有词典查找语料库中包含的具有代表性的石油专业术语 150 个, 制作语义关系训练数据集, 三类关系分别提取 1500 个词对. 每类关系取 1000 个词对作为训练集, 500 个词对作为测试集. 本文模型及参数设置如表 2 所示.

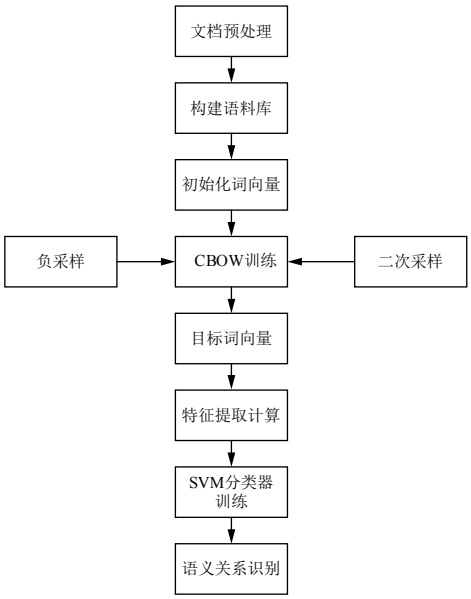

图 2 语义识别流程图

表 1 模型参数设置表

类别	实验设置
模型	CBOW
语料库	维基百科, 100 篇石油文档
向量维数	50, 100, 200, 400
上下文窗口大小	根据中心词所在句子自适应
是否采用 softmax	采用层次 softmax
负例数量	450
高频词截断阈值	500

4.2 评价指标

为了准确定义评价指标, 定义一个评价矩阵 $C(2\times3)$, $c_{ij}(i \in \{1,2,3\}, j \in \{0,1\})$表示词对数量, i 表示关系类别 (包括上下义关系、总分关系和类义关系), j=0 表示关系识别错误, j=1 表示关系识别正确. 本文使用准确率、召回率和 F 值对本文提出的方法进行评价. 其中, 准确率是语义关系识别正确的词语数与具有语义关系词语总数的比率, 衡量的是语义关系识别结果的查准率; 召回率是语义关系识别正确的词语数与实际具有语义关系词语总数的比率, 衡量的是语义关系识别结果的查全率. 两者取值在 0 和 1 之间, 数值越接近 1, 准确率或召回率就越高, 定义如下:

$$准确率(P) = \frac{c_{i1}}{c_{i0} + c_{i1}} \times 100\% \quad (11)$$

$$召回率(R)=\frac{c_{i1}}{\sum_{i=1,2,3} c_{i1}}\times 100\% \quad (12)$$

$$F=\frac{2RP}{R+P} \quad (13)$$

4.3 实验结果与分析

基于 4.1 节所述数据集, 利用第 3 节所述方法进行试验, 最后在测试集上进行评价.

表 2 给出了三种关系分别在 50 维、100 维、200 维和 400 维的词向量表示条件下, 基于向量差、向量和、向量和乘和向量连接这 4 中不同分类特征时 SVM 分类的 F 值.

表 2 不同词向量维度、不同特征下语义关系识别 F 值

关系类别	维度	f_{diff}	f_{add}	f_{mul}	f_{cat}
上下义关系	50	0.647	0.743	0.593	0.617
	100	0.715	0.675	0.644	0.737
	200	0.708	0.784	0.689	0.695
	400	0.794	0.827	0.762	0.792
总分关系	50	0.554	0.716	0.696	0.561
	100	0.715	0.723	0.744	0.602
	200	0.711	0.724	0.672	0.590
	400	0.789	0.746	0.793	0.683
类义关系	50	0.764	0.614	0.592	0.676
	100	0.717	0.729	0.720	0.717
	200	0.776	0.754	0.759	0.731
	400	0.761	0.829	0.797	0.804

为了更直观的观察不同维度对语义关系识别的效果, 将表 2 中的数据以折线图的形式进行展示, 如图 3、图 4、图 5 所示.

从表 2 和图 3、图 4、图 5 中可以看出, 不同词向量维度和向量特征对词汇语义关系的识别有一定的差异. 总体来看, 当词向量维度越大时, 识别效果越好. 向量和对上下义关系识别较好, 而向量乘则对总分关系识别较好. 综合识别效果和计算复杂度等因素, 最终选择维度为 200 的词向量进行特征组合试验, 测试不同特征对识别效果的影响, 实验结果如表 3 和图 6 所示.

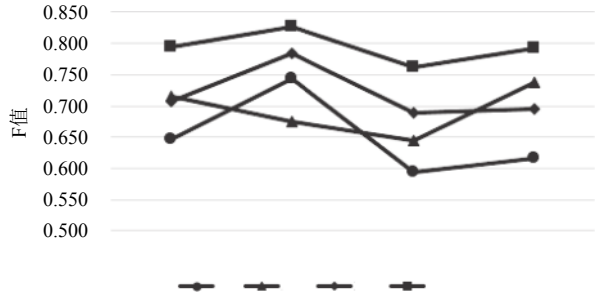

图 3 不同词向量维度下的上下义关系 F 值

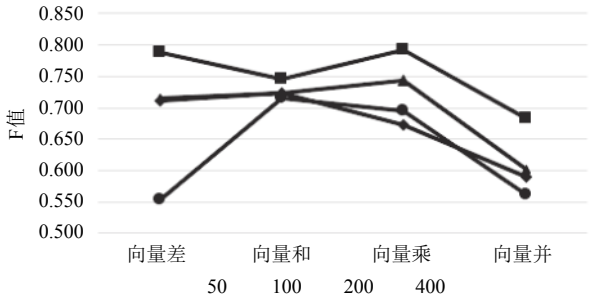

图 4 不同词向量维度下的总分关系 F 值

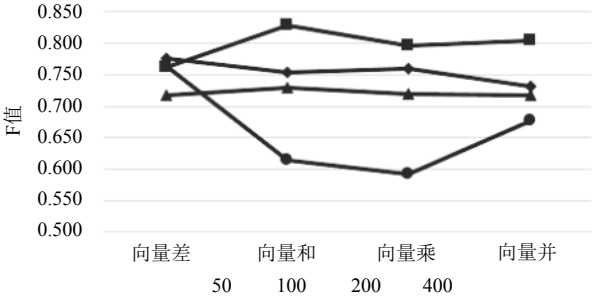

图 5 不同词向量维度下的类义关系 F 值

表 3 不同特征组合时的三种语义关系识别 F 值

	上下义关系 F 值	总分关系 F 值	类义关系 F 值
f_{diff}	0.708	0.711	0.776
f_{add}	0.754	0.724	0.754
f_{mul}	0.689	0.672	0.759
f_{cat}	0.695	0.59	0.731
$f_{diff}+f_{add}$	0.761	0.732	0.781
$f_{add}+f_{mul}$	0.712	0.722	0.782
$f_{add}+f_{cat}$	0.715	0.716	0.779
$f_{diff}+f_{add}+f_{mul}$	0.772	0.742	0.79
$f_{diff}+f_{add}+f_{cat}$	0.776	0.738	0.787
$f_{diff}+f_{add}+f_{mul}+f_{cat}$	0.791	0.784	0.802

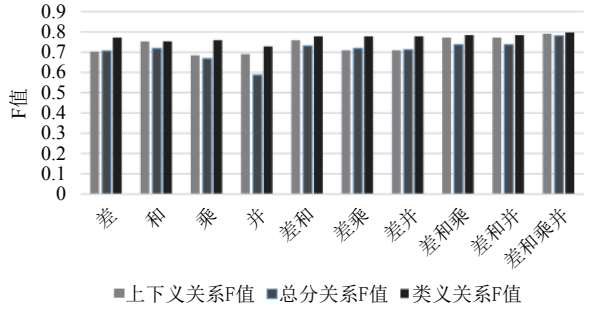

图 6 不同特征下三种关系 F 值

实验结果表明, 在使用单个特征进行识别时, f_{add}特征的综合性最好. 因此在f_{add}的基础上再组合其

他特征. 通过表 3 和图 6 可以看出, 随着特征的增加, 识别效果的 F 值也在增加, 说明增加特征可以提高语义关系的识别效果. 当四种特征全部组合时, 识别效果达到最好, 三种语义关系的 F 值分别达到了 0.791、0.784 和 0.802, 效果非常好.

最后与传统 CBOW 模型训练的词向量进行试验对比, 维度为 200 维, 特征包含向量差、和、乘、连接四种特征, 实验结果如表 4 和图 7 所示.

表 4 传统 CBOW 模型词向量与改进词向量 F 值

	上下义关系 F 值	总分关系 F 值	类义关系 F 值
传统 CBOW 模型词向量	0.655	0.701	0.687
改进的词向量	0.791	0.784	0.802

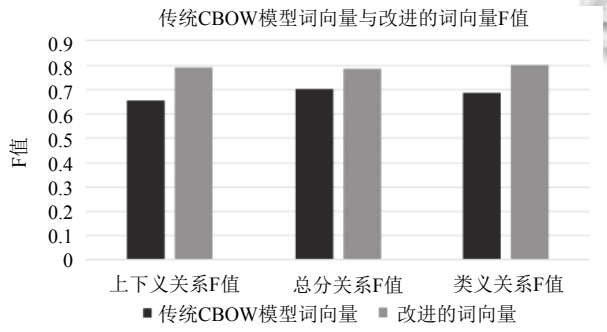

图 7 不同传统 CBOW 模型词向量与改进词向量 F 值

实验结果表明, 传统 CBOW 模型词向量对上下义关系、总分关系和类义关系识别的 F 值分别为 0.655、0.701 和 0.687, 改进的词向量对三种关系的识别 F 值分别为 0.791、0.784 和 0.802. 改进的词向量相对于传统的词向量, 对三种语义关系的识别效果提高了 20.8%、11.8% 和 16.7%, 识别效果提高显著.

5 总结与展望

石油领域文档的语义关系识别是构建石油领域本体的重要步骤. 本文提出基于改进词向量的石油文档语义关系识别算法, 利用 CBOW 模型对是由专业术语进行扩展训练, 增加正样本权重, 结合负采样和二次采样提高准确率和效率, 再利用 SVM 对语义关系进行识别. 实验结果表明, 本文提出的方法在石油领域专业术语语义关系识别方面具有明显的优势. 结合“向量差”、“向量和”、“向量乘”和“向量并”特征, 在语义关系识别方面具有很好的效果. 同时, 各种特征及其组合特征在测试集上的 F 值大部分都在 0.8 以下, 说明语义关系识别仍然不够准确. 在下一步工作中, 寻求更好的分类特征和新的词向量表示形式是研究重点.

参考文献

1 秦春秀, 祝婷, 赵捧未, 等. 自然语言语义分析研究进展. 图书情报工作, 2014, 58(22): 130–137.
2 Zhong MS, Hu Y, Liu L. Research on Chinese text segmentation based on quantified conceptual relations extracted from Chinese dictionary. Computer Engineering and Applications, 2008, 44(21): 25–29, 88.
3 Kurematsu M, Iwade T, Nakaya N, *et al*. DODDLE II: A domain ontology development environment using a MRD and text corpus. IEICE Transactions on Information and Systems, 2004, E87-D(4): 908–916.
4 易绵竹, 姚爱钢. 一种基于语义词典的俄语文本自动语义分析技术 SemLP. 中文信息处理前沿进展——中国中文信息学会二十五周年学术会议论文集. 北京, 中国. 2006.
5 Shamsfard M. Lexico-syntactic and semantic patterns for extracting knowledge from persian texts. International Journal on Computer Science and Engineering, 2010, 2(6): 2190–2196.
6 刘秀磊. 基于词法分析和语义分析的本体集成研究[博士论文]. 北京: 北京邮电大学, 2012.
7 孟晖, 王树良, 李德毅. 基于云变换的概念提取及概念层次构建方法. 吉林大学学报 (工学版), 2010, 40(3): 782–787.
8 Rios-Alvarado AB, Lopez-Arevalo I, Sosa-Sosa VJ. Learning concept hierarchies from textual resources for ontologies construction. Expert Systems with Applications, 2013, 40(15): 5907–5915. [doi: 10.1016/j.eswa.2013.05.005]
9 聂志强. 本体自动抽取中的概念相似性分析. 计算机工程与应用, 2007, 43(26): 159–163.
10 Ciaramita M, Gangemi A, Ratsch E, *et al*. Unsupervised learning of semantic relations between concepts of a molecular biology ontology. Proceedings. of the 19th International Joint Conference on Artificial Intelligence. Edinburgh, Scotland. 2005. 659–664.
11 刘宏哲. 文本语义相似度计算方法研究[博士论文]. 北京交通大学, 2012.
12 周慧霞. 基于词向量的中文词汇蕴涵知识获取研究[硕士论文]. 兰州: 西北师范大学, 2016.
13 蒋振超, 李丽双, 黄德根. 基于词语关系的词向量模型. 中文信息学报, 2017, 31(3): 25–31.
14 杜漫, 徐学可, 杜慧, 等. 面向情绪分类的情绪词向量学习. 山东大学学报 (理学版), 2017, 52(7): 52–58. [doi: 10.6040/j.issn.1671-9352.1.2016.072]
15 Collobert R, Weston J. A unified architecture for natural language processing: Deep neural networks with multitask learning. Proceedings of the 25th International Conference on Machine Learning. Helsinki, Finland. 2008. 160–167.
16 Li X, Qin T, Yang J, *et al*. LightRNN: Memory and computation-efficient recurrent neural networks. arXiv:1610.09893, 2016.
17 Bengio Y, Ducharme R, Vincent P. A neural probabilistic language model. Innovations in Machine Learning. Springer Berlin Heidelberg, 2006: 137–186.