

基于 BILSTM_CRF 的知识图谱实体抽取方法

翟社平^{1 2} 段宏宇¹ 李兆兆¹

¹(西安邮电大学计算机学院 陕西 西安 710121)

²(陕西省网络数据分析与智能处理重点实验室 陕西 西安 710121)

摘 要 针对传统知识图谱实体抽取方法需要大量人工特征和专家知识的问题,提出一种基于 BILSTM_CRF 模型的神经网络结构实体抽取方法。它既能使用双向长短时记忆网络 BILSTM(Bidirectional Long Short-Term Memory)提取文本信息的特征,又可利用条件随机场 CRF(Conditional Random Fields)衡量序列标注的联系。该方法对输入的文本进行建模,把句子中的每个词转换为词向量;利用 BILSTM 处理分布式向量得到句子特征;使用 CRF 标注并抽取实体,得到最终结果。实验结果表明,该方法的准确率和召回率更高,F1 值提升约 8%,具有更强的适用性。

关键词 知识图谱 实体抽取 神经网络 词向量 BILSTM_CRF 模型

中图分类号 TP391 文献标识码 A DOI: 10.3969/j.issn.1000-386x.2019.05.046

KNOWLEDGE GRAPH ENTITY EXTRACTION BASED ON BILSTM_CRF

Zhai Sheping^{1 2} Duan Hongyu¹ Li Zhaozhao¹

¹(School of Computer Science and Technology, Xi'an University of Posts and Telecommunications, Xi'an 710121, Shaanxi, China)

²(Shaanxi Key Laboratory of Network Data Analysis and Intelligent Processing, Xi'an 710121, Shaanxi, China)

Abstract In order to solve the problem that traditional knowledge atlas entity extraction method needed a lot of artificial features and expert knowledge, we proposed a neural network entity extraction method based on BILSTM_CRF model. It could use bidirectional long short-term memory(BILSTM) to extract the features of text information, and use conditional random fields(CRF) to measure the association of sequence labeling. In this method, the input text was modeled, and every word in the sentence was transformed into a word embedding. The distributed vectors were processed by BILSTM to get sentence features. And the CRF tagging and entity extraction were used to get the final results. The experimental results show that the proposed method has higher accuracy and recall rate, and the value of F1 is increased by about 8%, which has better applicability.

Keywords Knowledge graph Entity extraction Neural network Word embedding BILSTM_CRF model

0 引 言

知识图谱是一种包括众多实体及其关系的语义网络,实现了语义表述和智能连接,是人工智能领域的重要构成技术。目前,知识图谱已经在电商搜索、知识库问答以及知识管理等领域得到深层应用^[1]。实体抽

取,也称命名实体学习或命名实体识别,是知识图谱构建中最为重要和关键的部分,实现了从文本数据中抽取实体信息^[2]。高精度的实体抽取很大程度上保障了所构建知识图谱的高适用性。

早期使用基于规则的方法实现实体抽取,通过领域专家和语言学者手工制定有效规则,将文本与规则进行匹配识别出实体^[3]。比如,在中文系统中,人名的

收稿日期:2018-10-01。工业和信息化部通信软科学项目(2018-R-26);陕西省自然科学基金资助项目(2012JM8044);陕西省社会科学基金资助项目(2016N008);陕西省教育厅科学研究计划资助项目(12JK0733);西安邮电大学研究生创新基金项目(CXL2016-13)。翟社平,副教授,主研领域:语义网、知识图谱。段宏宇,硕士生。李兆兆,硕士生。

下文用词可以是“吃”、“同学”等,而单位组织名的尾部可以是“集团”、“学校”等,利用这些规则,可以将人名、地点名、组织名等实体抽取出来。传统采用制定规则的实体识别方法有: Grishman 开发的 Proteus 系统^[4]、Black 开发的 FACILE 系统^[5]等。然而,基于规则的方法构建过程中往往需要大量的语言学知识,而且费时费力、可移植性不好。之后,知识图谱实体抽取任务被部分机器学习算法解决。机器学习使用大量的实验语料训练学习,得到最终的标注模型,完成对实体的词性标注识别。常应用到实体识别任务中的模型包括隐马尔科夫模型 HMM^[6]、支持向量机模型 SVM^[7]、条件随机场模型 CRF^[8]等。目前比较流行的是条件随机场模型,通过人工定义特征模板,利用上下文标注信息发掘命名实体。这种利用人工特征的方法取得了很好的识别效果,但人工制定特征代价较昂贵,不同领域的特征不尽相同,导致识别方法不通用,工作量大。伴随着计算机硬件能力的提高、文本单词向量化表示的广泛应用,神经网络模型在处理实体抽取任务方面表现出更高的效率,例如循环神经网络^[9]、卷积神经网络^[10]以及长短时记忆网络^[11]。使用神经网络提取文本特征不再依赖人工特征和领域知识,减少了人工特征提取所需代价,有效地提高了系统效率。

本文对中文领域的命名实体进行识别,结合双向长短时记忆网络以及条件随机场模型,自动提取句子特征,提高实体抽取效率。

1 相关工作

1.1 词向量

在传统的实体抽取方法中,单词采用 0 和 1 组成的稀疏向量表示,向量中某一分量的值为 1,其余分量的值为 0,1 的位置表示该词在词典中的编号^[12]。这种方法可以完成一部分实体抽取任务,但存在缺点,一是词典规模通常比较大,维度高,容易造成维度灾难和数据稀疏问题;二是采用稀疏向量,导致前后单词彼此独立,无法得到两个词的关联信息。

为了计算词本身之间的语义相似度,深度学习利用低维、稠密的实值向量标识单词,称词向量。词向量包含了文本的语义信息,采用欧式距离或余弦夹角等方法来衡量两个向量词之间的距离,发掘两个词的关系。Bengio^[13]在 2003 年提出了一种基于深度学习的语言模型,如图 1 所示。

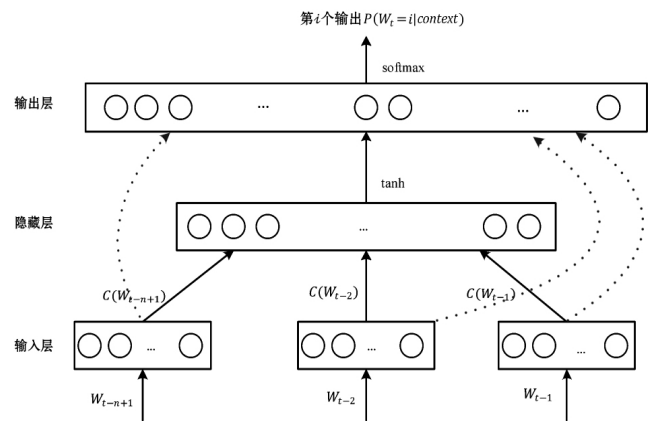


图 1 神经网络语言模型

该方法引入词向量来构建概率模型,公式表示如下:

$$p(s) = p(w) = p(w_1, w_2, \dots, w_T) = \prod_{t=1}^T p(w_t | \text{context}) \quad (1)$$

context 为单词 w_t 的上下文,通过深度学习方法来实现概率语言模型。类似于传统 n-gram 概率模型,利用前 $n-1$ 个词 $w_{t-n+1}, \dots, w_{t-2}, w_{t-1}$ 来推测下一个词 w_t 出现的概率。输入向量 x 由前 $n-1$ 个词 $w_{t-n+1}, \dots, w_{t-2}, w_{t-1}$ 的词向量连接组成,中间隐藏层引入 tanh 函数去线性化,输出层获得隐藏层的结果后利用 Softmax 函数归一化处理,得到概率值 $P(w_t | \text{context})$,最后应用极大似然估计进行训练,得到单词 w_t 的词向量 $C(w_t)$ 。

除了上述用于训练词向量的模型方法外,目前常用 Google 开源的词向量训练框架 Word2vec 生成词向量^[14]。Word2vec 采用向量相似度表示某种隐含的语义关系,使用向量空间的向量运算简化文本内容处理。

1.2 BILSTM 神经网络

实体抽取是典型的文本序列标注问题,RNN 网络能够有效地利用数据的序列信息,并具有一定的记忆功能,是一种有效地解决序列标注任务的神经网络,但其无法很好地处理长距离依赖问题,LSTM 引入门限制机制对历史信息进行过滤,有效地解决了这个问题。由于 LSTM 只是利用当前词的上文信息,而在实体抽取处理中,对当前词的识别同样需要下文信息,双向长短时记忆网络模型 BILSTM 结构能同时包含文本上下文信息^[15],原理如图 2 所示。在 BILSTM 模型中,包含前向和后向两个 LSTM 层,连接着同一个输出层。由于训练序列的前向和后向 LSTM 层组成了 BILSTM 网络结构,所以在实体抽取中,BILSTM 模型能兼顾上下文信息,自动提取句子特征,获得更好的结果。

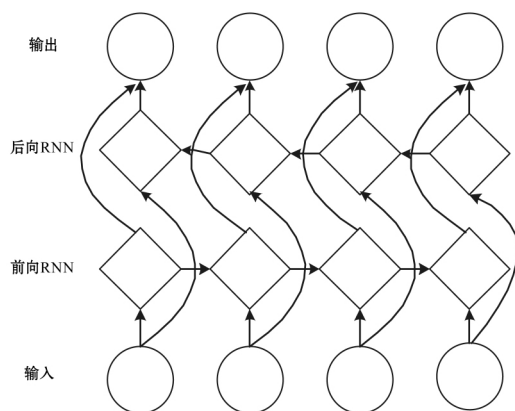


图2 双向长短时记忆网络

2 基于 BILSTM-CRF 的实体抽取

传统实体抽取方法 CRF 模型需要大量的人工特征, 现有的 BLSTM 神经网络模型对每个单词标注的过程是独立的分类, 不能直接利用上文标签信息, 导致预测出的标签序列可能是错误的, 影响实体抽取效果。本文用 BILSTM-CRF 模型完成实体抽取任务, 通过 BILSTM 网络学习特征, 避免人工提取特征的复杂性, 利用 CRF 层考虑前后句子的标签信息, 使实体标注抽取不再是对每个单词独立的分类, 实体抽取更为准确、高效。

2.1 模型架构

本文提出的实体抽取 BILSTM-CRF 模型分为三层, 首先是输入层, 其次是隐含层, 最后是标注层, 模型架构如图3所示。

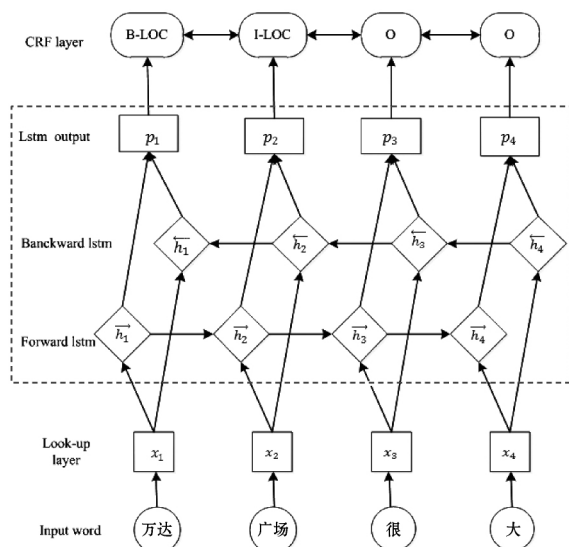


图3 BILSTM-CRF 模型架构

look-up 是输入层, 主要负责将窗口的词进行向量化映射, 使用训练好的词向量矩阵将输入窗口的每个词 x_i 映射为分布式向量 $x_i, x_i \in R^d, d$ 是向量的维度。为防止出现过拟合状况, 设置 dropout 参数。

BILSTM 是隐含层, 实现自动获得句子特征。将 look-up 输出的字向量序列 (x_1, x_2, \dots, x_n) 作为双向 LSTM 网络每个时间点的初始输入值。在 BILSTM 网络层中, 输入向量序列的顺序序列作为前向 LSTM 层的输入, 逆序序列则作为后向 LSTM 层的输入。在 t 时刻, 模型将正向 LSTM 输出的隐状态序列 $(\vec{h}_1, \vec{h}_2, \dots, \vec{h}_n)$ 与反向 LSTM 输出的隐状态序列 $(\overleftarrow{h}_1, \overleftarrow{h}_2, \dots, \overleftarrow{h}_n)$ 进行按位置拼接 $h_t = [\vec{h}_t; \overleftarrow{h}_t] \in R^m$, 得到完整的隐状态序列 $(h_1, h_2, \dots, h_n) \in R^{n \times m}, m$ 是隐状态向量维度。为了自动提取句子特征, 在隐含层中接入一个线性变换层, 把之前得到的隐状态序列从 m 维映射到 k 维, k 代表标注集中所有的标签数, 将结果记作矩阵 $P = (p_1, p_2, \dots, p_n) \in R^{n \times k}$, 把 $p_i \in R^k$ 的每一位 P_{ij} 项都视作将字 x_i 分类到第 j 个标签的打分值。这时对输出结果进行归一化处理, 表示只对每个位置独立分类, 并没有使用上下文已标记过的信息。

CRF 是标注层, 也叫逻辑回归层, 进行语句序列标注。CRF 层引入一个状态转移矩阵 M , 矩阵 M 的每个元素 M_{ij} 表示从 i 变化到 j 的可能性, 实现利用此前标注过的信息对一个新的位置进行标注。如果记一个标签序列为 $y = (y_1, y_2, \dots, y_n)$, 对于句子 x , 模型预测标签等于 y 的打分为 $\text{score}(x, y) = \sum_{i=1}^n (M_{y_{i-1}y_i} + P_{iy_i})$, 可以看出打分结果由两部分组成, 一部分分值是由 LSTM 的输出 p_i 决定, 另外一部分则取决于 CRF 的变化矩阵 M , 最终使用 Softmax 函数进行归一化处理, 得到结果 $P(y|x) = \frac{\exp(\text{score}(x, y))}{\sum_{y'} \exp(\text{score}(x, y'))}$ 。

2.2 模型训练

2.2.1 前向神经网络

模型通过前向神经网络将输入的特征向量经过层层推导得到最后的输出, 并通过这些输出解决实体抽取任务。本实验实体抽取模型的输入层是大小为 $n \cdot V$ 的窗口词向量, n 代表窗口的规模, V 指词向量的维度。向量化的字符作为隐含层的输入数据, 隐含层得到的结果则是标注层的特征。标注层表示的是输入层的字符为每个分类的可能性。网络架构的前向神经网络函数如下:

$$h = W(x_1, x_2) + b^1 \quad (2)$$

$$a = f(h) \quad (3)$$

$$z = g(U^T a + b^2) \quad (4)$$

式中: 向量 (x_1, x_2) 是输入层的词向量组, h 代表输入词向量通过线性变换得到的特征向量, a 为隐含变换层的输出结果, f 是去线性化的激活函数, 采用双曲正

切函数 $\frac{d}{dx} \tanh x = 1 - \tanh^2 x$ 输出区间在 $(-1, 1)$ 之间, 整个函数是以 0 为中心。 f 和 g 指的是网络模型函数, W 、 U 代表模型设计参数 b 是用来去线性化的标量值, z 表示最后分类的概率值。

2.2.2 损失函数

在前向神经网络的逻辑回归层, 常使用 Softmax 函数计算一种状态序列 $y \in Y_x$ 的概率。针对单个样本的损失函数, 应用极大似然估计来定义:

$$l(x, y) = -\log(p(y|x)) \quad (5)$$

模型进行训练时, 这种基于单样本梯度参数更新方法不仅造成模型训练速度慢的问题, 还容易带来优化波动现象, 本文引入批量样本的梯度进行参数更新, 定义模型的损失函数为:

$$L(X, Y) = -\log \prod_{(x, y) \in (X, Y)} p(y|x) = \sum_{(x, y) \in (X, Y)} l(x, y) \quad (6)$$

式中: X 、 Y 为批量训练样本, 大小为 batch。

2.2.3 参数优化

参数优化是神经网络模型训练的核心任务, 利用反向传播算法可以计算损失函数对每个参数的梯度, 并结合学习率更新参数, 从而提高训练模型的质量。传统梯度下降法每次都采用相同的学习速率 η 来更新所有的参数 θ_i , 即:

$$g_{t+1} = \nabla_{\theta} L(\theta) \quad (7)$$

$$\theta_{t+1} = \theta_t - \eta \cdot g_{t+1} \quad (8)$$

这种采用固定学习速率的方法难以适用于稀疏梯度, 而 Adagrad 方法可以实现为各个参数动态地分配不同的学习率。在利用 Adagrad 方法进行模型参数的更新时, 对样本较少的特征进行较大幅度的参数调整, 对频繁的特征进行微调, 通过这种方式实现对学习率进行一定的约束, 其形式化定义公式如下:

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{G_{t+1}} + \varphi} \cdot g_{t+1} \quad (9)$$

式中: φ 用来防止分母为零项, $G_t \in \mathbf{R}^m$ 是一个对角矩阵, 矩阵中的元素 $G_{t, ii}$ 表示每个参数 θ_i 从最初状态到 t 时刻所有梯度平方的求和。在 t 时刻, 学习速率在对应参数 θ_i 的前一时刻梯度基础上被更改, 观察式(9)发现, 当 θ_i 的梯度 g_i 持续较小时, 对参数进行大幅度调动, 而参数已经接近最优, 仅仅需要微调。

2.2.4 训练算法

图 4 展示了利用反向传播算法优化网络模型的过程。反向传播是一个重复反馈的过程。首先采用较小单位训练数据, 通过前向传播算法得到神经网络的预测结果, 然后利用损失函数计算预测值和真实值之间

的差距, 最后用反向传播更新模型参数的取值, 模型训练算法如算法 1 所示。

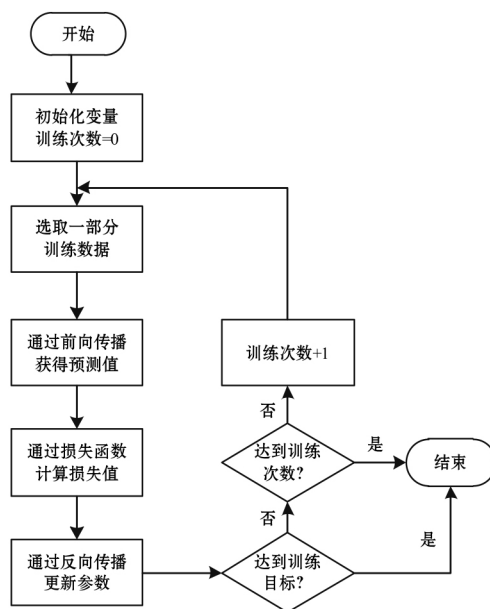


图 4 神经网络反向优化流程

算法 1 BILSTM_CRF 模型训练

输入: 模型参数 θ , 损失函数 $L(X, Y)$, 训练样本 X , 训练样本 Y

输出:

1. 对于每个 epoch 循环
2. 对于每个 batch 循环
3. Function forward($\theta, L(X, Y), X, Y$)
4. BILSTM 前向训练
5. CRF 前向训练
Calculate the $L(X, Y)$ and
 $g_{t+1} = \nabla_{\theta} L(\theta)$ /* 计算损失函数 */
6. End function
7. Function backward(θ, g_{t+1})
8. $\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{G_{t+1}} + \varphi} \cdot g_{t+1}$ /* 参数更新 */
9. CRF 反向传播及更新参数
10. BILSTM 反向传播及更新参数
11. End function
12. 结束 batch 循环
13. 结束 epoch 循环
14. Return θ

3 实验及分析

3.1 实验准备

3.1.1 环境准备

为了解决 CPU 训练模型需要消耗较长时间的问题, 本实验用 NVIDIA GeForce GTX 1060 显卡对模型进行训练, 使用该型号 GPU 训练时间缩短为小时级

别,实验需要配置的软硬件如表1所示。

表1 软硬件配置表

	项目	环境
硬件	GPU	GTX 1060
	内存	8 GB
	硬盘	500 GB
软件	操作系统	Ubuntu 16.04
	Python 版本	Python 2.7
	TensorFlow 版本	TensorFlow 1.7

3.1.2 语料准备

实验使用2014年1月人民日报语料库样本,其中训练语料占实验样本的80%,剩余作为实验测试语料,分别对样本中的人物姓名(PER)、地点名(LOC)和组织机构名(ORG)进行识别,具体的数据统计如表2所示。

表2 语料数据统计表

语料	训练集	测试集	所有实体
人名	7 100	1 776	8 876
地名	2 870	717	3 587
组织名	3 540	987	4 437

监督训练方式主要标注模型包括 BIO、BIEO、BIE-SO 等,本文利用的模式是 BIO,B 代表命名实体首字,I 代表实体中间,O 代表非实体。因此,B-PER、I-PER 分别表示人名首字、人名非首字,B-LOC、I-LOC 表示地点名首字、地点名非首字,B-ORG、I-ORG 指组织名首字、组织名非首字,O 代表非命名实体。

3.1.3 词向量

实体标注抽取实验需要一个初始词向量,相对于直接随机初始化词向量,使用大量非标注样本语料训练词向量效率更高。本实验词向量训练工具用谷歌开源 word2vec 语言模型 CBOW。设置 CBOW 模型参数:词向量的维度大小为 100、200 和 300 维;滑动窗口大小为 5;训练迭代次数 100;设置学习速率 0.05 等。

3.1.4 神经网络

谷歌开源的 TensorFlow 计算框架能够很好地支持深度学习的各种算法,本文利用 TensorFlow 框架来训练 BILSTM-CRF 实体抽取模型,具体的参数设置为:神经网络层数 2;隐含层节点数 300;最大梯度值 5;学习速率 0.02;过拟合参数 0.6;批量样本大小 128 等。

3.2 评价指标

本实验使用准确率 P(Precision)、召回率又叫查全率 R(Recall)以及 F1 值对模型的性能进行评价^[16]。3 个评价指标的公式定义如下:

$$P = \frac{\text{正确识别的实体个数}}{\text{实体总数}} \quad (10)$$

$$R = \frac{\text{正确识别的实体个数}}{\text{识别出的实体总数}} \quad (11)$$

$$F1 = \frac{2 \times P \times R}{P + R} \quad (12)$$

3.3 实验结果与分析

(1) 实验采用传统条件随机场 CRF 模型、双向长短时记忆网络 BILSTM 模型,混合方法 LSTM-CRF 模型作为对比实验。传统的 CRF 方法利用实验语料人工设计特征模板,而不是用词向量表示字符,模板设计如图5所示。

```

1 # Unigram
2 U00:%x[-2,0]
3 U01:%x[-1,0]
4 U02:%x[0,0]
5 U03:%x[1,0]
6 U04:%x[2,0]
7 U05:%x[-1,0]/%x[0,0]
8 U06:%x[0,0]/%x[1,0]
9 ...
10 U10:%x[-2,1]
11 U11:%x[-1,1]
12 U12:%x[0,1]
13 U13:%x[1,1]
14 U14:%x[2,1]
15 U15:%x[-2,1]/%x[-1,1]
16 U16:%x[-1,1]/%x[0,1]
17 U17:%x[0,1]/%x[1,1]
18 U18:%x[1,1]/%x[2,1]
19 ...
20 U20:%x[-2,1]/%x[-1,1]/%x[0,1]
21 U21:%x[-1,1]/%x[0,1]/%x[1,1]
22 U22:%x[0,1]/%x[1,1]/%x[2,1]
23
24 # Bigram
25 B

```

图5 CRF 特征模板

每个模板特征用 % [row,col] 标记实验数据,row,col 表示距离目标数据的行、列长度。“U01: % x [-1,0]”指基于当前观测 X_t ,考虑前一个观测 X_{t-1} 的第 $X_{t-1}^{(0)}$ 个特征以及当前状态 Y_t 。条件随机场方法根据这些特征模板和标注标签训练生成特征选择模型,接着对测试语料的词汇进行识别,计算出词汇为某一类命名实体的概率。通过实验,四种模型的性能对比如表3所示。

表3 不同模型对比表

模型类别	实体总数	P/%	R/%	F1/%
CRF	16 900	89.50	75.88	82.13
BILSTM	16 900	88.21	74.23	80.62
LSTM-CRF	16 900	90.58	80.34	85.15
BILSTM-CRF	16 900	94.15	86.63	90.23

BILSTM 在所有测试模型中 F1 值最低,神经网络模型结构较为复杂,需要大量的训练语料,而本实验语料库较小,模型得不到充分训练,性能受到影响;其次,BILSTM 对建模的状态是相互独立的,不能有效利用上

下已经标注过的序列信息,影响抽取效率。

混合模型 LSTM_CRF 比其中两种单一模型 $F1$ 值高,识别效果较好,但因 LSTM 网络只能提取序列标注模型的上文信息特征,无法利用下文信息特征,抽取模型未达到最好效果。

观察表 3 可以发现,相比于其他三种模型方法,基于 BILSTM_CRF 模型的实体抽取方法有更高的准确性和查全率,其中 $F1$ 值相比传统 CRF 模型提高约 8%,表明基于 BILSTM_CRF 的方法较其他方法具有较大优势。BILSTM_CRF 结合预训练向量,使用 BILSTM 选取最优特征工程,利用 CRF 句子级别的标签信息优化 BILSTM 结构,结合两者获得最佳的识别效果。

(2) 通过 word2vec 工具生成不同维度的向量对 BILSTM_CRF 模型的实体抽取效率也有影响,实验训练 100、200 以及 300 维度的向量分别进行对比,对于每一组实验,均循环 5 次并求平均值,结果如图 6 所示。

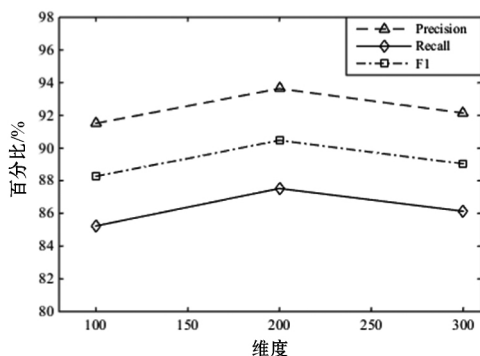


图 6 不同维度词向量实验对比

由图 6 可以发现,实体抽取实验所得的准确率、召回率以及 $F1$ 值在词嵌入向量维度为 200 最高,模型获得最佳性能。维度过低,获取的特征不完整,出现欠拟合情况;维度太高,语料中的噪声容易被捕获,产生过拟合状况。

(3) 在 BILSTM_CRF 模型中,网络结构复杂度和模型误差大小受隐含层节点数目的影响。实验训练包含不同隐含层节点数量的模型作对比,每种参数设置都执行 5 次,计算求得 $F1$ 值平均值。图 7 说明了 $F1$ 均值随着隐含层节点数目的变化。

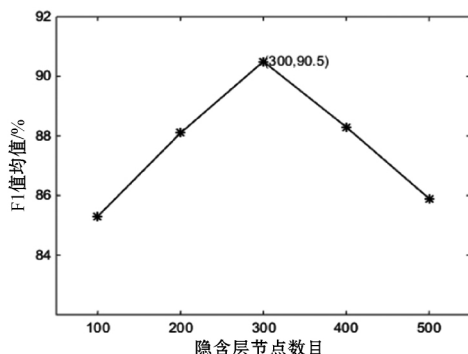


图 7 $F1$ 值均值随隐含层节点数目变化

由图 7 可知, $F1$ 值均值随着节点数量的增加而提高,在节点数量为 300 时达到一个最高值,随后又呈现出递减趋势。在隐含层节点数目是 300 时,模型获得最好实体抽取效果。分析原因是隐含层节点太少时,模型得不到充分训练,性能较差;节点数过多,又容易造成模型过度拟合和浪费训练时间等问题,影响实验效率。

4 结 语

针对目前知识图谱实体抽取方法依赖大量人工特征和领域专家的问题,提出了一种结合双向长短时记忆网络和条件随机场模型的实体抽取方法。该方法融入了词语的上下文信息,把词的分布式表示引入到特征提取,充分利用词语标签的前后联系,识别效果得到进一步提高。实验结果表明,基于 BILSTM_CRF 模型的实体抽取方法与传统方法相比, $F1$ 测度值更高,有更好的实体识别效果。

本文实体抽取模型训练需要大量标注语料,但在一些领域并没有海量的标注数据,所以基于所提方法引入迁移学习,使用少量的标注数据进行实体抽取将是以后研究的重点。

参 考 文 献

- [1] 杜泽宇,杨燕,贺樛. 基于中文知识图谱的电商领域问答系统[J]. 计算机应用与软件, 2017, 34(5): 153-159.
- [2] 翟社平,郭琳,高山,等. 一种采用贝叶斯推理的知识图谱补全方法[J]. 小型微型计算机系统, 2018, 39(5): 995-999.
- [3] 张海楠,伍大勇,刘悦,等. 基于深度神经网络的中文命名实体识别[J]. 中文信息学报, 2017, 31(4): 28-35.
- [4] Grishman R. The NYU system for MU-C6 or where's the syntax? [C]//Message Understanding Conference, Columbia, Maryland, USA, 2014: 13-31.
- [5] Brin S. Extracting Patterns and Relations from the World Wide Web [M]//The World Wide Web and Databases. Springer Berlin Heidelberg, 2016: 172-183.
- [6] 韩霞,黄德根. 基于半监督隐马尔科夫模型的汉语词性标注研究[J]. 小型微型计算机系统, 2015, 36(12): 2813-2816.
- [7] 张小龙,刘书妍,刘满华. 基于级联支持向量机融合多特征的人脸检测[J]. 计算机应用与软件, 2016, 33(4): 151-154, 207.
- [8] 何炎祥,罗楚威,胡彬尧. 基于 CRF 和规则相结合的地理命名实体识别方法[J]. 计算机应用与软件, 2015, 32(1): 179-185.

(下转第 280 页)

5 结 语

智能手机具有多品牌竞争扩散的特点。而目前,我们尚未见到具备强大销量预测能力的多品牌产品扩散模型出现。消费者的购买决策受产品、消费者及其所处社会环境中众多因素的影响,找到能将这些因素有机结合的一套建模理论和方法,以确保所构建的模型能准确刻画消费者的决策逻辑及其对产品竞争与扩散过程的影响逻辑,是构建具备强大销量预测能力的多品牌产品扩散模型的关键。本文运用多 Agent 理论和模糊推理技术,综合产品、消费者和社会环境中的重要因素所构建的多品牌产品扩散模型,被大量实验证明能较好地预测多品牌智能手机的销量走势。本文的研究不仅丰富和发展了创新扩散理论,而且还可为相关企业的产品营销提供决策参考。

参 考 文 献

- [1] 艾兴政,唐小我. 广告媒介下两种产品竞争与扩散模型研究[J]. 管理工程学报,2000,14(3): 19-22.
 - [2] 谭建,王先甲. 替代性产品在广告媒介下的扩散动态研究[J]. 软科学,2014,28(3): 110-113,118.
 - [3] Erickson G M. An oligopoly model of dynamic advertising competition [J]. European Journal of Operational Research, 2009,197(1): 374-388.
 - [4] Eliashberg J, Jeuland A P. The impact of competitive entry in a developing market upon dynamic pricing strategies [J]. Marketing Science, 1986,5(1): 20-36.
 - [5] 丁士海,韩之俊. 考虑竞争与重复购买因素的耐用品牌扩散模型[J]. 系统工程理论与实践,2011,31(7): 1320-1327.
 - [6] Guseo R, Mortarino C. Within-brand and cross-brand word-of-mouth for sequential multi-innovation diffusions [J]. IMA Journal of Management Mathematics, 2014,25(3): 287-311.
 - [7] Muller E, Peres R. The effect of social networks structure on innovation performance: A review and directions for research [J]. International Journal of Research in Marketing, 2018.
 - [8] Kiesling E, Gunther M, Stummer C, et al. Agent-based simulation of innovation diffusion: a review [J]. Central European Journal of Operations Research, 2012,20(2): 183-230.
 - [9] Schramm M E, Trainor K J, Shanker M, et al. An agent-based diffusion model with consumer and brand agents [J]. Decision Support Systems, 2010,50(1): 234-242.
 - [10] Kim S, Lee K, Cho J K, et al. Agent-based diffusion model for an automobile market with fuzzy TOPSIS-based product adoption process [J]. Expert Systems with Applications, 2011,38(6): 7270-7276.
 - [11] 李英,胡剑. 基于智能体的多类新能源汽车市场扩散模型[J]. 系统管理学报,2014,23(5): 711-716.
 - [12] Stummer C, Kiesling E, Günther M, et al. Innovation diffusion of repeat purchase products in a competitive market: An agent-based simulation approach [J]. European Journal of Operational Research, 2015,245(1): 157-167.
 - [13] Jiang G, Tadikamalla P R, Shang J, et al. Impacts of knowledge on online brand success: an agent-based model for online market share enhancement [J]. European Journal of Operational Research, 2016,248(3): 1093-1103.
 - [14] Mamdani E H, Assilian S. An experiment in linguistic synthesis with a fuzzy logic controller [J]. International Journal of Man-Machine Studies, 1975,7(1): 1-13.
 - [15] Barabási A, Albert R. Emergence of Scaling in Random Networks [J]. Science, 1999,286(5439): 509.
- ~~~~~
- (上接第 268 页)
- [14] 李玮,张大方,黄昆,等. 面向大数据处理的高精度多维计数布鲁姆过滤器[J]. 电子学报,2015,43(4): 652-657.
 - [15] Li K, Zhu Y, Yang J, et al. Video super-resolution using an adaptive superpixel-guided auto-regressive model [J]. Pattern Recognition, 2016,51(C): 59-71.
 - [16] Agarwal A, Maheswaran R, Kurths J, et al. Wavelet Spectrum and Self-Organizing Maps-Based Approach for Hydrologic Regionalization—a Case Study in the Western United States [J]. Water Resources Management, 2016,30(12): 4399-4413.
- ~~~~~
- (上接第 274 页)
- [9] Zhang J. RNN-BLSTM Based Multi-Pitch Estimation [C]// INTERSPEECH Germany: Inter-speech 2016: 1785-1789.
 - [10] 冯多,林政,付鹏,等. 基于卷积神经网络的中文微博情感分类[J]. 计算机应用与软件,2017,34(4): 157-164.
 - [11] Graves A, Mohamed A R, Hinton G. Speech recognition with deep recurrent neural networks [C]//IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2013: 6645-6649.
 - [12] 谢志宁. 中文命名实体识别算法研究[D]. 杭州: 浙江大学,2017.
 - [13] Bengio Y, Duchme R. A neural probabilistic language model [J]. Journal of Machine Learning Research, 2003,3(6): 1137-1155.
 - [14] 王蕾. 基于神经网络的中文命名实体识别研究[D]. 南京: 南京师范大学,2017.
 - [15] 冯艳红,于红,孙庚,等. 基于 BLSTM 的命名实体识别方法[J]. 计算机科学,2018,45(2): 261-268.
 - [16] 李航. 统计学习方法[M]. 北京: 清华大学出版社,2012.