

文章编号:1008-1542(2020)03-0225-08

开放科学(资源服务)标识码(OSID):



基于 BERT-BiGRU-Attention 的在线健康社区 用户意图识别方法

迟海洋^{1,2}, 严馨^{1,2}, 周枫^{1,2}, 徐广义³, 张磊^{1,2}

(1.昆明理工大学信息工程与自动化学院,云南昆明 650500;2.昆明理工大学云南省人工智能重点实验室,云南昆明 650500;3.云南南天电子信息产业股份有限公司,云南昆明 650040)

摘要:针对传统用户意图识别主要使用基于模板匹配或人工特征集合方法导致成本高、扩展性低的问题,提出了一种基于 BERT 词向量和 BiGRU-Attention 的混合神经网络意图识别模型。首先使用 BERT 预训练的词向量作为输入,通过 BiGRU 对问句进行特征提取,再引入 Attention 机制提取对句子含义有重要影响力的词的信息以及分配相应的权重,获得融合了词级权重的句子向量,并输入到 softmax 分类器,实现意图分类。爬取语料实验结果表明,BERT-BiGRU-Attention 方法性能均优于传统的模板匹配、SVM 和目前效果较好的 CNN-LSTM 深度学习组合模型。提出的新方法能有效提升意图识别模型的性能,提高在线健康信息服务质量、为在线健康社区问答系统提供技术支持。

关键词:自然语言处理;意图识别;在线健康社区;BERT 词向量;BiGRU;Attention 机制

中图分类号:TP391 **文献标识码:**A **doi:**10.7535/hbkd.2020yx03004

An online health community user intention identification method based on BERT-BiGRU-Attention

CHI Haiyang^{1,2}, YAN Xin^{1,2}, ZHOU Feng^{1,2}, XU Guangyi³, ZHANG Lei^{1,2}

(1.Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, Yunnan 650500, China; 2.Yunnan Key Laboratory of Artificial Intelligence, Kunming University of Science and Technology, Kunming, Yunnan 650500, China; 3. Yunnan Nantian Electronic Information Industry Company Limited, Kunming, Yunnan 650040, China)

Abstract: Aiming at the problem of high cost and low expansibility of traditional user intention recognition, which mainly uses template matching or artificial feature set, a hybrid neural network intention recognition model based on BERT word embedding and BiGRU-Attention was proposed. First, the word embedding pre-trained by BERT was used as the input, and

收稿日期:2020-04-17;修回日期:2020-05-23;责任编辑:冯民

基金项目:国家自然科学基金(61562049,61462055)

第一作者简介:迟海洋(1994—),男,山东日照人,硕士研究生,主要从事自然语言处理方面的研究。

通讯作者:严馨副教授。E-mail:kg_yanxin@sina.com

迟海洋,严馨,周枫,等.基于 BERT-BiGRU-Attention 的在线健康社区用户意图识别方法[J].河北科技大学学报,2020,41(3):225-232.

CHI Haiyang, YAN Xin, ZHOU Feng, et al. An online health community user intention identification method based on BERT-BiGRU-Attention[J]. Journal of Hebei University of Science and Technology, 2020, 41(3): 225-232.

the features of the interrogative sentences were extracted by BiGRU. Then, the attention mechanism was introduced to extract the information of words that have important influence on the meaning of sentences and allocate the corresponding weights, so as to obtain the sentence embedding that integrates the word-level weights and input it into the softmax classifier to realize intention classification. According to the experiment on the crawling corpus, it shows that the performance of BERT-BiGRU-Attention method is better than that of traditional template matching, SVM and lately popular CNN-LSTM deep learning combined model. The proposed method can effectively improve the performance of intention recognition model and the quality of online health information service, which provide technical support for the online health community question answering system.

Keywords: natural language processing; intention identification; online health communities; BERT word embedding; bidirectional gated recurrent unit(BiGRU); Attention mechanism

信息技术正在影响传统医疗领域,它提供了更低的成本和更好的医疗服务。近年来,在线健康社区盛行,并且发展迅速。在线健康社区为用户提供了一种简单方便的方式来寻求医疗健康信息服务,用户可以在其中发布描述自己健康状况或查询需求。大多数在线健康社区支持用户和临床医生之间的交互,同时可以向合适的临床医生实时推荐新生成的查询,以加快响应过程。在这个过程中,在线健康社区产生了大量有价值的医疗问答数据、记录,包含了大量的医学、健康信息和知识。这些问答数据可以用于许多医学相关任务,如知识抽取、问答系统、疾病推断等。如果能够利用这些信息提前挖掘出用户咨询意图,那么可以进一步向用户提出合理建议以及推荐合适的医师、科室。

意图识别旨在了解用户潜在需求,可以看成分类问题。事先在该领域范围内定义可能的意图类别,再用分类方法将问句划分到合适的类中。目前在意图识别研究中,大致分为3种方法。

一是基于规则、模板匹配的方法,该方法需要通过专家人工提取针对特定问题类型的特征规则来确定问题所属类型。如 RAMANAND 等^[1]提出基于规则和图的方法来获取意图模板进而实现用户的消费意图。RAY 等^[2]提出了8种问题模式并结合 Wikipedia 与 WordNet 将问题分成7个大类。LIU 等^[3]通过匹配核心元组来分类。LI 等^[4]发现不同表达方式会导致规则模板数量的增加。这种方法虽不需训练大量数据,但费时费力、可扩展性弱。

二是基于统计的机器学习方法,该方法通过对标注语料进行统计学习,把分类模型的选择和分类特征的提取作为重点关注对象。常用的方法有支持向量机(SVM)、朴素贝叶斯(NB)和逻辑回归等。陈浩辰^[5]分别使用 SVM 和 NB 分类器实现了微博消费意图的分类。贾德华^[6]优化了 SVM 参数,得到了一种 Ada-Boost-POSVM 分类模型。SILVA 等^[7]使用特征组合方式对问题分类。LIU 等^[8]提出将词性、词袋和句法依存树结合,并通过计算核函数值的方法探究问题结构。这种方法需要人工设定特征,成本较高,当数据集变化时,会演变成特征设计、特征选取问题,且特征准确性得不到保障,最终难以准确理解用户文本的深层次语义信息。

三是基于深度学习的方法,该方法通过自我学习方式学习到句子中的内在语义和句法特征^[9]。常见的方法有卷积神经网络(CNN)、循环神经网络(RNN)、长短时记忆网络(LSTM)、门控循环单元(GRU)、注意力(Attention)机制等。

1) 基于卷积神经网络的识别:KIM^[10]将 CNN 用于文本分类。HASHEMI 等^[11]使用 CNN 提取文本向量表示作为查询分类特征来识别用户查询意图。该方法减少了大量的特征工程,但只能提取意图文本的局部语义特征。

2) 基于循环神经网络及其变体的意图识别:BHARGAVA 等^[12]将上下文信息应用到意图识别任务中,提高了意图识别正确率。RAVURI 等^[13]提出将 LSTM 与 RNN 两种模型应用于意图分类问题,并比较这两种方法,发现 LSTM 比 RNN 具有更好的建模能力和记忆功能。GRU 是 LSTM 的一种改进^[14],双向门控循环单元(BiGRU)可以对意图文本进行更好的特征表示。RAVURI 等^[15]通过实验发现 GRU 与 LSTM 模型在意图分类任务上性能几乎一致,GRU 模型更简单,参数更少。

3) 基于深度学习模型组合运用的意图识别:综合考虑各种深度学习的优缺点,近年来大部分研究将不同优势的深度学习模型进行组合进而更好地挖掘用户意图。钱岳等^[16]提出基于 Convolutional-LSTM 模型进行出行消费意图的识别,提高了 F 值。余慧等^[17]提出了基于短文本主题模型(biterm topic model, BTM)和双向门控循环单元(bidirectional gated recurrent unit, BiGRU)的意图识别模型(BTM-BiGRU),在

用户就医意图识别上获得很好的效果。黄佳伟^[18]提出了 Character-CNN-BGRU 的深度学习组合模型,验证了组合模型在任务上的优势。该模型结构复杂、训练时间较长。

4) 基于 Attention 机制模型的意图识别:LIN 等^[19]提出了一种引入自注意力机制提取句子表示的模型,可以获取句子的多种语义信息,有助于意图识别的研究。CAI 等^[20]提出了一种 CNN-LSTM 并引入 Attention 机制的方法来理解在线健康社区的用户查询意图,取得了不错的效果。

针对在线健康社区的特点并结合深度学习方法以及 BERT 在 NLP 领域的绝对优势,本文提出了一种基于 BERT 词向量和 BiGRU-Attention 的混合神经网络意图识别模型。

1 分类体系

分类体系是意图识别的依据,其设计直接影响意图识别的准确性。问题分类可以表示为一种映射函数: $G: X \rightarrow C$, 其中 X 表示问题集合, $C = \{c_1, c_2, \dots, c_n\}$ 是由 n 个问题类别组成的分类体系, $\forall x \in X, G$ 负责利用先验知识将其映射到分类体系中某一个类别 c_i 中。目前分类体系尚无统一的标准,具体到在线健康医疗领域,若使用开放领域的问句分类难以满足其分类需求,还需要面向领域知识的主题分类。本文在健康问答社区的^[21]主题分析以及综合分析健康问答网站(<http://www.wenda120.com>)中提出的问题基础上,根据期望查询目标将用户问句大致归为几个类别,如表 1 所示。

表 1 问题类别及示例
Tab.1 Question categories and examples

问题类别	问题示例
解释类	什么是阿尔兹海默症?
症状类	阿尔兹海默症有什么症状?
诊断类	如何诊断阿尔兹海默症?
咨询类	患有阿尔兹海默症会对生活有什么影响? 有康复的可能吗?
治疗类	如何治疗阿尔兹海默症?
病因类	阿尔兹海默症发病原因是什么?
检查类	糖尿病要做什么检查?
药品类	哪些药物可以治疗糖尿病?
择医类	糖尿病属于哪个科室? 哪家医院、大夫治疗好?
并发症类	糖尿病有哪些并发症?
病情管理类	患有糖尿病,平时应该注意些什么?
健康生活方式类	晨炼时,适合做什么运动?

2 模型构建

本文提出的基于 BERT 词向量和 BiGRU-Attention 的意图识别模型如图 1 所示,由输入层、BERT 词向量层、BiGRU 层、词级别 Attention 机制层和分类层组成。

2.1 BERT 词向量层

语义消歧决定了语言处理任务的性能,是自然语言处理任务的一个核心和难点,模型在处理训练集时应当考虑和联系句子上下文内容。ELMo 和 GPT 等模型都取得了不错的效果,而 DEVLIN 等^[22]在此基础上改进了语言模型单向信息流的问题,使用多层 Transformer^[23]结构。Transformer 相比 LSTM 模型没有长度限制问题,具备更好地捕获上下文信息特征的能力,通过注意力机制将任意位置的 2 个单词的距离转换成 1,有效地解决了 NLP

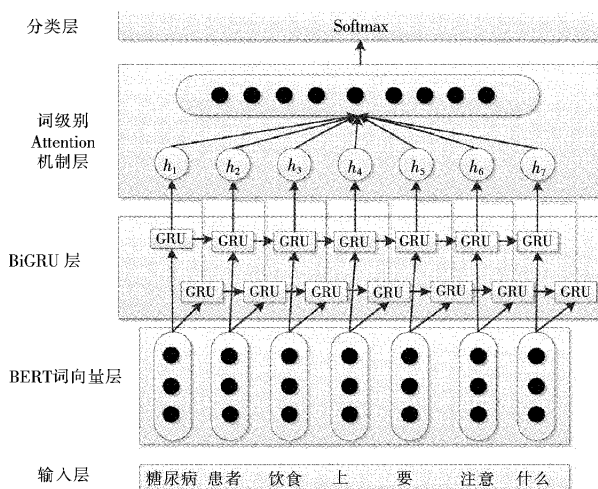


图 1 基于 BERT 词向量和 BiGRU-Attention 的意图识别模型结构

Fig.1 Intention recognition model structure based on BERT word embedding and BiGRU-Attention

中棘手的长期依赖问题。相比单向训练模式, BERT双向训练模型捕获上下文信息会更加全面, 因此本文使用 BERT 生成输入层句子的词向量。

2.2 BiGRU 层

门控循环单元(GRU)是 LSTM 的一个变体, 通过“门”的结构让信息选择性通过, 在保证学习效果的同时又使结构更加简单, 提高了训练速度。GRU 由重置门 r_t 和更新门 z_t 组成, 其中更新门由输入门和遗忘门组成。其工作流程如下:

首先, 由重置门控制并决定从旧元胞状态和当前输入中丢弃哪些信息, 如式(1)所示。

$$r_t = \sigma(W_r x_t + U_r h_{t-1}), \quad (1)$$

式中: σ 代表 Sigmoid 激活函数; x_t 代表当前输入; h_{t-1} 代表上一时刻隐层的输出。

接下来, 决定将哪些新信息保存到元胞状态, 具体分为 2 部分。

1) 由更新门实现忘记之前信息和添加信息的操作, 如式(2)所示。

$$z_t = \sigma(W_z x_t + U_z h_{t-1}). \quad (2)$$

2) 由 tanh 层创建一个新的候选值 \tilde{h}_t , 该值可能会加入到元胞状态中, 如式(3)所示。

$$\tilde{h}_t = \tanh(W_h x_t + r_t * U_h h_{t-1}). \quad (3)$$

最后, 把这 2 个组合起来用于更新旧元胞状态 h_{t-1} 到新元胞状态 h_t , 如式(4)所示。

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t. \quad (4)$$

单向 GRU 忽略了未来上下文的信息, 而改进的双向 GRU 可用同一个训练序列分别向前和向后各训练一个 GRU 模型, 再将 2 个模型的输出进行线性组合, 使得序列中每一个节点都能完整地依赖于所有上下文信息, 故对于本文任务, 选用能更好理解用户意图的双向 GRU。

2.3 词级别 Attention 机制层

为捕获更直接的语义依赖关系, 本文将 BiGRU 模型的全连接输出输入到 Attention 机制层中, 采用 Attention 机制对用户提问语句进行编码。并不是所有的词对句子意义的表达都有同等的作用^[24], 如用户输入“如何治疗糖尿病?”, 用户意图是寻求治疗, 属于“治疗类”, 在分类过程中“治疗”一词对正确分类贡献最大, 所占权重应该最高, 因此采用词级别 Attention 机制^[24-25]来提取对句子含义重要的词。

给定一个序列 $S = (w_1, w_2, \dots, w_T)$, T 表示序列长度, 在第 i 时刻输入的第 t 个单词的词向量为 c_{it} , 通过双向 GRU 后可以充分学习上下文之间关系, 进行语义编码, 具体过程如式(5)所示。

$$h_{it} = \text{BiGRU}(c_{it}), \quad t \in [1, T]. \quad (5)$$

词级别 Attention 机制可通过以下 3 个步骤实现。

1) 使用多层感知机获取 h_{it} 的隐藏表示 (u_{it}), 如式(6)所示。

$$u_{it} = \tanh(w_w h_{it} + b_w), \quad (6)$$

式中: w_w 为模型权重; b_w 为偏置。

2) 计算 u_{it} 和词级别上下文 u_w 的相似性并将其作为权衡单词重要性指标, 由 softmax 函数计算归一化权重 α_{it} , α_{it} 与每一时刻的输入状态和 u_w 相关, 如式(7)所示。

$$\alpha_{it} = \frac{\exp(u_{it}^T u_w)}{\sum_t \exp(u_{it}^T u_w)}. \quad (7)$$

其中单词上下文向量 u_w 可以看作是对输入的一种语义表示, 是在训练过程中随机初始化和共同学习的。

3) 计算句子向量, 见式(8)。

$$S = \sum_{i=1}^n \alpha_{it} h_{it}. \quad (8)$$

该向量中包含了各个输入状态的重要程度信息。

2.4 分类层

该层主要任务是构建识别用户意图的分类器, 获取用户问题语句的语义表示对于每一个分类标签的得分向量, 并输出最终意图分类标签。本文采用 softmax 分类器输出各类别的概率, 计算方法见式(9)。

$$y = \text{softmax}(w_f s + b_f). \quad (9)$$

本文通过反向传播机制对模型中的参数不断进行训练和更新,使用交叉熵损失函数对模型进行优化,见式(10)。

$$\text{loss} = - \sum_{i=1}^D \sum_{j=1}^C \hat{y}_i^j \log(y_i^j), \quad (10)$$

式中: D 为训练数据集大小; C 为意图类别数; \hat{y}_i 为句子的真实意图类别标签。

3 实验与分析

3.1 实验数据及评价指标

本文所使用的实验语料,是爬取了寻医问药网(<http://z.xywy.com>)、好大夫在线(<https://www.haodf.com>)等国内知名在线健康医疗网站近一年的用户问题语句,共计 78 573 条,涉及内容丰富。首先将爬取语料均分为 4 部分;其次由 12 名标注人员 3 人 1 组分别对爬取内容中的某一部分进行意图类别标注,共计 4 组;再通过匹配每组人员的语料标注结果,去掉标注不一致部分;最终将各组标注一致的数据合并,共获取 67 930 条记录。其意图类别标注结果及实验语料例句如表 2 所示。

表 2 意图类别标注结果
Tab.2 Annotated results of the intention category

意图类别	记录条数	实验语料例句
解释类	5 191	晚上睡觉时,腿突然很猛地抖了一下,这是怎么回事?
症状类	4 725	2 个月的宝宝得了肺炎会有什么症状和表现?
诊断类	9 053	大便出血,会得直肠癌吗?
咨询类	10 612	孩子做了脊柱畸形手术,还能长高吗?
治疗类	12 857	声带小结要怎么治疗?
病因类	7 078	手上发白是什么原因造成的?
检查类	3 066	心脏病患者需要去医院做什么检查?
药品类	4 259	我没有感冒,就是喉咙痛、说话沙哑,吃什么药好啊?
择医类	2 668	天津医院男科哪里看的好?
并发症类	1 284	糖尿病会引发什么其他的疾病和症状?
病情管理类	3 436	地中海贫血在饮食上应该注意什么?
健康生活方式类	3 057	一日三餐的食用量分别多少比较好?
其他	644	3 岁孩子左脚痛走不了,手脚有瘀痕。

在线健康社区中的用户意图识别属于多分类问题,因此采用准确率 P 、召回率 R 及 F_1 值对每种类别进行评价,见式(11)一式(13)。

$$P = TP / (TP + FP), \quad (11)$$

$$R = TP / (TP + FN), \quad (12)$$

$$F_1 = 2PR / (R + P), \quad (13)$$

式中: TP 表示属于类 A 且被正确分类到类 A 中的样本数; FN 表示属于类 A 却被错误分类到其他类别中的样本数; FN 表示不属于类 A 且被正确分类到其他类别中的样本数; FP 表示不属于类 A 却被错误分类到类 A 中的样本数。

3.2 实验参数设置

不同参数的组合与设置会对实验产生不同的效果,本文基于 Pytorch 深度学习框架,神经网络初始化使用 Kaiming 方法,损失函数使用交叉熵损失函数,模型训练优化使用 Adam 方法。初始学习率设为 0.001,学习率衰减速率设为 0.000 1,词向量维度设为 200, $\text{batch_size}=64$, $\text{Epoch}=10$, $\text{attention_size}=128$,在输入层 Dropout 设为 0.2,隐藏层 Dropout 设为 0.5,其余权重、偏置等参数随模型优化不断变化。

3.3 实验结果与分析

在对比实验中,将数据集按 9:1 分为训练集和测试集。首先在数据预处理过程中去除杂质文本,再使用 Jieba 分词工具对用户聊天文本分词,去除停用词、标点符号、特殊字符等,最后使用谷歌开源提供的 BERT 词向量进行训练以量化文本。

为验证本文提出的方法效果性,除基于 BERT-BiGRU 和 Word2vec-BiGRU-Attention 混合模型外,还与传统的模板匹配、SVM 以及目前效果较好的深度学习组合模型 CNN-LSTM^[16] 作对比,不同方法在测试语料上整体分类性能对比实验结果如表 3 所示,不同方法在其各个意图分类中 F_1 值的比较如表 4 所示。

表 3 不同方法对比实验结果

Tab.3 Comparative results of different experimental methods

实验模型	P/%	R/%	F_1 /%
模板匹配	77.3	69.7	73.3
SVM	79.9	71.7	75.6
CNN-LSTM	82.6	77.0	79.7
BERT-BiGRU	85.4	82.3	83.8
Word2vec-BiGRU-Attention	87.1	83.8	85.4
BERT-BiGRU-Attention	89.6	86.3	87.9

表 4 不同方法在其各个意图分类中 F_1 值比较

Tab.4 Comparison of F_1 values in each intention classification in different methods

意图类别	F_1					
	模板匹配	SVM	CNN-LSTM	BERT-BiGRU	Word2vec-BiGRU-Attention	BERT-BiGRU-Attention
解释类	78.3	77.1	83.6	86.5	90.7	91.6
症状类	74.6	75.3	78.4	82.0	79.5	87.1
诊断类	70.4	74.6	80.8	85.0	87.8	88.7
咨询类	68.9	71.7	82.1	85.6	84.7	86.6
治疗类	76.5	78.9	85.3	88.2	91.3	92.9
病因类	78.8	77.9	90.9	90.7	92.0	92.3
检查类	71.2	72.7	77.7	78.6	83.4	87.2
药品类	74.7	75.4	76.2	80.7	85.2	88.6
择医类	72.5	75.5	80.5	83.8	84.0	86.1
并发症类	80.2	82.8	76.9	84.0	85.6	86.4
病情管理类	67.6	70.6	68.9	80.6	78.2	82.8
健康生活方式类	66.2	72.3	74.7	79.3	81.8	84.4

由表 3、表 4 可知,本文提出的基于 BERT 词向量和 BiGRU-Attention 的混合神经网络意图识别模型在准确率、召回率及 F_1 值的表现性能均优于其他 5 种方法。

传统模板匹配和 SVM 方法的意图识别效果最差。究其原因在于基于规则的模板匹配覆盖率低、模板和特征抽取有局限性,而传统机器学习方法基于特征工程,以词或句法结构作为分类特征难以获取句子的深层语义信息,识别度较低。

混合模型 CNN-LSTM 相对传统方法有明显的提升。CNN 可以抽取到句子丰富的局部特征,与具有可以学习序列信息的 LSTM 结合,优势互补,更加充分提取了句子的特征,从而实现了分类性能的提升。本文选取收敛速度更快、参数较少、计算简单的 BiGRU 模型,可以在保证准确率的同时提高效率。

对比基于 Word2vec, BERT 方法训练的词向量, BERT 预训练的词向量更有优势,在多个任务中表现优异,解决了单向信息流问题,能够充分捕获句子特征信息,进而提高分类性能。

对比 BiGRU 和 BiGRU-Attention 模型,引入 Attention 机制性能显著提升,通过 Attention 机制层可以对句子含义表达有突出贡献的词分配更多权重、减小无关信息干扰,提高了意图识别准确度。Attention 权重示例如图 2 所示。

通过对比可以得出,本文提出的方法在在线健康社区用户意图识别任务上具有良好的表现。

4 结 语

本文提出了一种基于BERT词向量和BiGRU-Attention的在线健康社区用户意图识别方法,首先利用BERT预训练词向量的优势挖掘用户问句信息,其次使用结构简单、高效的双向门控循环单元进行完整上下文学习,最后采用词级别的注意力机制来增强问句表示和意图识别效果。基于健康社区查询意图的实验表明,与传统模板匹配、统计机器学习以及现有的较好深度学习方法相比,本文提出的方法能有效提高意图识别的准确率和召回率,更好地理解用户意图。

本文不足之处是未能考虑用户问句自身的特征信息,在后续工作中,将考虑对用户问句的一些特征信息进行提取和挖掘,如使用主题模型对语义进行挖掘,同时,尝试进一步优化神经网络模型以及有效融合特征信息,以提高用户意图识别性能。

参考文献/References:

- [1] RAMANAND J, BHAVSA R K, PEDANEKA R N. Wishful thinking: Finding suggestions and ‘buy’ wishes from product reviews [C]//Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text. Stroudsburg: Association for Computational Linguistics, 2010: 54-61.
- [2] RAY S K, SINGH S, JOSHI B P. A semantic approach for question classification using WordNet and Wikipedia [J]. Pattern Recognition Letters, 2010, 31(13): 1935-1943.
- [3] LIU X M, LIU L. Question classification based on focus [C]//Proceeding of the 2012 International Conference on Communication Systems and Network Technologies. Washington DC: IEEE Computer Society, 2012: 512-516.
- [4] LI X, DAN R. Learning question classifiers: The role of semantic information [J]. Natural Language Engineering, 2015, 12(3): 229-249.
- [5] 陈浩辰. 基于微博的消费意图挖掘[D]. 哈尔滨: 哈尔滨工业大学, 2014.
CHEN Haochen. Consumption Intention Mining Based on Microblog [D]. Harbin: Harbin Institute of Technology, 2014.
- [6] 贾俊华. 一种基于 AdaBoost 和 SVM 的短文本分类模型[D]. 天津: 河北工业大学, 2016.
JIA Junhua. A Short Text Classification Model Based on Combination of AdaBoost and SVM [D]. Tianjin: Hebei University of Technology, 2016.
- [7] SILVA J, COHEUR L, MENDEZ A C, et al. From symbolic to sub-symbolic information in question classification [J]. Artificial Intelligence Review, 2011, 35(2): 137-154.
- [8] LIU L, YU Z, GUO J, et al. Chinese question classification based on question property kernel [J]. International Journal of Machine Learning and Cybernetics, 2014, 5(5): 713-720.
- [9] 李超, 柴玉梅, 南晓斐, 等. 基于深度学习的问题分类方法研究 [J]. 计算机科学, 2016, 43(12): 115-119.
LI Chao, CHAI Yumei, NAN Xiaofei, et al. Research on problem classification method based on deep learning [J]. Computer Science, 2016, 43(12): 115-119.
- [10] KIM Y. Convolutional neural networks for sentence classification [C]//Proc of the 2014 Conference on Empirical Methods in Natural Language Processing. [S.l.]: [s.n.], 2014: 1746-1751.
- [11] HASHEMI H B, ASIAEE A, KRAFT R. Query intent detection using convolutional neural networks [C]//International Conference on Web Search and Data Mining, Workshop on Query Understanding. [S.l.]: [s.n.], 2016. doi: 10.1145/1235.
- [12] BHARGAVA A, CELIKYILMAZ A, HAKKANITUR D, et al. Easy contextual intent prediction and slot detection [C]//IEEE International Conference on Acoustics. [S.l.]: [s.n.], 2013: 8337-8341.
- [13] RAVURI S V, STOLCKE A. Recurrent neural network and LSTM models for lexical utterance classification [C]//16th Annual Conference of the International Speech Communication Association. [S.l.]: [s.n.], 2015: 135-139.
- [14] DEY R, SALEMT F M. Gate-variants of gated recurrent unit (GRU) neural networks [C]//IEEE 60th International Midwest Symposium on Circuits and Systems. [S.l.]: [s.n.], 2017: 1597-1600.
- [15] RAVURI S V, STOLCKE A. A comparative study of recurrent neural network models for lexical domain classification [C]//Proc of the 41st IEEE International Conference on Acoustics, Speech, and Signal Processing. [S.l.]: [s.n.], 2016: 6075-6079.
- [16] 钱岳, 丁效, 刘挺, 等. 聊天机器人中用户出行消费意图识别 [J]. 中国科学: 信息科学, 2017, 47(8): 997-100.

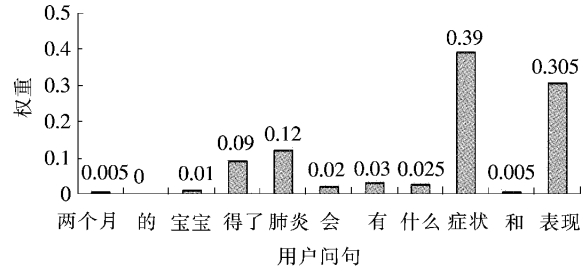


图2 数据集中一条用户问句的 Attention 权重示例

Fig.2 Attention based weighting sample for a sentence from the dataset

- QIAN Yue, DING Xiao, LIU Ting, et al. Identification method of the user's travel consumption intention in chatting robot [J]. Scientia Sinica Information, 2017, 47(8): 997-100.
- [17] 余慧,冯旭鹏,刘利军,等.聊天机器人中用户就医意图识别方法[J].计算机应用,2018,38(8):2170-2174.
YU Hui, FENG Xupeng, LIU Lijun, et al. Identification method of user's medical intention in chatting robot [J]. Journal of Computer Applications, 2018, 38(8): 2170-2174.
- [18] 黄佳伟.人机对话系统中用户意图分类方法研究[D].武汉:华中师范大学,2018.
HUANG Jiawei. Research on Intent Classification in Dialogue Systems [D]. Wuhan: Central China Normal University, 2018.
- [19] LIN Zhouhan, FENG Minwei, SANTOS C N D, et al. A Structured Self-attentive Sentence Embedding [EB/OL]. <https://arxiv.org/pdf/1703.03130.pdf>, 2020-03-15.
- [20] CAI R, ZHU B, JI L, et al. An CNN-LSTM attention approach to understanding user query intent from online health communities [C]// 2017 IEEE International Conference on Data Mining Workshops (ICDMW). [S.l.]: IEEE, 2017: 430-437.
- [21] 金碧漪,许鑫.网络健康社区中的主题特征研究[J].图书情报工作,2015(12):102-107.
JIN Biyi, XU Xin. Research on theme features in online health community [J]. Library and Information Service, 2015(12): 102-107.
- [22] DEVLIN J, CHANG M W, LEE K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding [C]// Proceedings of NAAACL-HLT. [S.l.]: [s.n.], 2019: 4171-4186.
- [23] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]:// Proc of the 31st Conf on Neural Information Processing Systems. New York: Curran Associates, 2017: 5998-6008.
- [24] 张志昌,张珍文,张治满.基于IndRNN-Attention的用户意图分类[J].计算机研究与发展,2019,56(7):1517-1524.
ZHANG Zhichang, ZHANG Zhenwen, ZHANG Zhiman. User intent classification based on IndRNN-Attention [J]. Journal of Computer Research and Development, 2019, 56(7): 1517-1524.
- [25] YANG Z, YANG D, DYER C, et al. Hierarchical attention networks for document classification [C]// Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. [S.l.]: [s.n.], 2016: 1480-1489.

向本期载文的审稿专家致谢

本期《河北科技大学学报》共发表论文10篇。这些论文的发表是与有关专家的认真审读、细查资料、推敲分析、中肯评价分不开的。对此,本编辑部特向这些专家表示敬意,对他们的辛勤劳动表示感谢。本期载文的审稿专家名单如下(按姓名的汉语拼音顺序排列):

柴争义 陈维林 董选普 封筠 高凯 金泰松 敬霖 李文斌
梁英 刘爱军 罗玉涛 唐连生 田丽梅 王英 王珍珍 邢书明
杨易 于道永 张宁 张永才

(本刊编辑部)