

智能催收系统中自然语言理解模块设计

赵 冬, 穆平安

(上海理工大学 光电信息与计算机工程学院, 上海 200093)

摘 要: 随着社会信贷消费的流行, 待催收的违约用户逐渐增多。针对此问题, 对 Rasa 对话系统开发框架中的自然语言理解模块进行改进, 显著提升智能催收系统中自然语言理解模块的准确率, 并对最终得到的训练模型进行性能评估。在所提出的算法模型中, 用户输入文本经面向特定领域的预训练语言模型处理后输入到意图识别与槽位填充模块。在提出的模型中采用递归卷积神经网络模型进行意图识别, 门控循环单元及条件随机场模型完成槽位填充任务, 最后与原始预训练语言模型及其他主流算法进行比较。意图识别、槽位填充模型实验结果显示, 本文方法 F1 值分别达 95.75%、95.88%, 均优于其他主流算法, 验证了所提出算法的可行性。

关键词: 智能催收系统; Rasa; 自然语言理解; 预训练语言模型

DOI: 10.11907/rjdk.212434

中图分类号: TP391

文献标识码: A

开放科学(资源服务)标识码(OSID):

文章编号: 1672-7800(2022)009-0026-07



Design of Natural Language Understanding Module in Intelligent Collection System

ZHAO Dong, MU Ping-an

(School of Optional Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: With the popularity of social credit consumption, the number of defaulting users to be collected is gradually increasing. To solve this problem, this paper improves the natural language understanding module in the development framework of Rasa dialogue system, significantly improves the accuracy of the natural language understanding module in the intelligent collection dialogue system, and evaluates the performance of the final training model. In the algorithm model proposed in this paper, the user input text is processed by the domain oriented pre training language model and input to the intention recognition and slot recognition module. In the model proposed in this paper, the recursive convolution neural network model is used for intention recognition, and the gated loop unit and conditional random field model are used to complete the task of slot recognition. Finally, by comparing with the original pre training language model and other mainstream algorithms, F1 values in intention recognition and slot recognition reach 95.75% and 95.88% respectively, which are better than other mainstream algorithms, which verifies the feasibility of the algorithm proposed in this paper.

Key Words: intelligent collection system; Rasa; natural language understanding; pre training language model

0 引言

在金融领域, 银行主要起着为企业或个人提供周转资金的杠杆作用, 银行的盈利主要来源于企业和个人的贷款业务。随着经济的快速发展, 截至 20 年末, 我国金融机构人民币各项贷款余额 172.75 万亿元, 同比增长 12.8%。但

是贷款贷出去存在有风险, 当个人或贷款企业无法偿还贷款时, 就形成了坏账, 随之而来的是待催收违约客户也逐渐增多^[1]。在银行内部建立催收部门或者打官司成本高且收效甚微, 于是不得不将不良账户送到第三方机构进行催收, 第三方往往采用电话轰炸、骚扰通讯录等极端暴力催收手段。催收过程中催收人员往往掌握着大量债务人的个人信息及隐私, 以此威胁债务人还款, 这使得债务人

收稿日期: 2021-10-19

基金项目: 2021 年学位点引导布局与建设培育项目(XWDB2021105)

作者简介: 赵冬(1997-), 男, 上海理工大学光电信息与计算机工程学院硕士研究生, 研究方向为深度学习、自然语言处理; 穆平安(1964-), 男, 上海理工大学光电信息与计算机工程学院教授、硕士生导师, 研究方向为图像处理、测试信息获取与处理、在线检测技术与装置。本文通讯作者: 穆平安。

还面临着信息被泄露的风险。随着国家对相关问题的重视,催收行业的监管也日趋严格,暴力催收不可取。得益于近年来海量增长的大数据和飞速发展的人工智能技术^[2],利用机器人进行催收取得了新的进展。催收过程中智能机器人的使用,不仅可以节省大量人工成本,还可以保护债务人的隐私,提升催收效率和客户满意度。智能催收机器人可以自动记录每一通电话的内容,以语音和文本的方式存储在企业账户后台,为企业后期数据分析、话术优化做准备,工作效率也是人工坐席的3~4倍。

目前在实际应用中,NLU模块面临的主要问题有:①训练深度模型需要大量的标注语料,尤其是面向特定领域的中文标注语料;②深度学习模型对语言的理解只停在语义的表层,在复杂任务下,对话系统结合已有知识进行逻辑推理的能力较弱,不能利用意图识别和槽位填充两者之间的联系进行系统联合建模;③用户对话自由度高,没有明确的规范句式,使得对话存在实时性、不规则性,从而影响了NLU模块的准确性。

为了解决以上问题,本文基于开源对话系统机器学习框架Rasa,在自然语言理解模块中加入最新提出的面向金融领域预训练语言模型(FinBERT),结合深度学习算法^[3]提升自然语言理解模中意图识别^[4]与槽位填充功能的准确率;并与原始预训练语言模型及其他深度学习算法比较,结合对话系统在智能催收领域的实际应用,对NLU模块的意图识别和槽位填充功能进行模型训练;最后结合Rasa框架内置的模型评估命令,对训练好的NLU模块进行测试与分析。

1 相关工作

1.1 语料库

在实际应用中,制约NLU模块发展的一个关键因素是相关模型需要大量的标注语料。在NLU领域,斯坦福大学自然语言处理小组在SQuAD1.0^[5]的基础上加入了“不可回答的问题”将数据集扩充为SQuAD2.0版本^[6]。使语料与真实场景更为接近。2018年,Reddy等^[7]发布了CoQA数据集,该数据集面向多轮对话,相比SQuAD数据集,CoQA种类更加丰富,因为每组对话都需要对上下文进行理解推断,所以更能反映人类真实对话场景。2020年,微软亚洲研究院发布了多轮对话推理数据集MuTual,在该数据集上,最先进的方法准确率只能达到71%,远低于人类94%的平均水平^[8]。

1.2 神经网络模型

神经网络是一种模仿动物神经网络行为特征,进行分布式并行信息处理的算法模型,这种网络通过调节内部大量节点间相互连接的关系,从而达到信息处理目的。目前,基于神经网络模型的各种算法已广泛应用于自然语言理解相关研究。

Hemphill等^[9]使用基于双向模型的RNN语义框架解析网络结构并考虑意图识别和槽位填充两个任务之间的交互作用,联合执行两个任务,实验表明,该方法在ATIS数据集上优于其他方法。Chen等^[10]提出基于注意力机制的双向长短期记忆网络(LSTM)模型,用作呼和浩特市公交查询库中的意图识别与槽位填充任务,实验表明,基于字向量的方法要优于基于词向量的方法,实验中提出的模型在F1值上优于其他原始LSTM方法。

2014年,KIM^[11]改进了卷积神经网络模型(CNN)算法并将其用于句子分类任务,结果表明,在情感分析和问题分类在内的4个任务上均有所改进。Xu等^[12]提出一种基于CNN的意图识别与槽位填充的联合检测模型,该神经网络模型可以看成是一种三角形条件的条件随机场(CRF)模型。该模型使用CNN提取特征,这也是使用神经网络模型在意图识别与槽位填充联合任务的尝试。Vu^[13]将改进的CNN模型应用于序列标注任务,模型保留了序列的上下文信息,该方法不需要先验的语言知识,性能优于之前的基于RNN模型。

1.3 联合模型

考虑到意图检测和时隙填充之间的密切关系,最新的研究中采用联合模型以利用任务间的共享知识。联合模型又分为显示联合建模和隐式联合建模,其中隐式联合模型指该模型采用共享编码器捕获共享特征,没有任何显式交互。Liu等^[14]使用了一种带有共享机制的编码器—解码器结构,共用于意图识别和槽位填充。Zhang等^[15]构造一个关键动词和其上下文信息相结合的模型以解决问题,进一步提高slot-gated模型性能。双向流相互作用功能意味着模型考虑了意图识别与槽位填充中的交叉影响,Wang等^[16]提出使用两个相关的双向LSTM并考虑跨意图识别和槽位填充交叉影响的模型体系结构。

1.4 预训练模型

BERT模型是Google于2018年发布的一种语言表示模型,一经发布就在多个自然语言处理任务中取得了出色的成绩。该模型通过在大量大规模无标注语料上的训练得到一个通用的“语言理解”模型,其输入由初始字向量、文本向量和位置向量组成,在下游的特定文本分析任务中无需对特定任务的体系结构进行大量修改,只需对参数进行微调即可^[17]。在基于BERT的模型中,每个话语以[CLS]开始,以[SEP]结束,其中[CLS]是表示整个序列的特殊符号,[SEP]是分隔非连续词块序列的特殊符号。此外,特殊词块[CLS]的表示用于意图检测,而其他词块表示用于槽位填充。Chen等^[18]具体研究了BERT模型在自然语言理解领域的应用,BERT模型用于提取上下文的嵌入信息以进行意图识别及槽位填充,在与未经预训练的语言模型相比,文中提到的方法取得了优异成绩。战保行^[19]采用BERT模型实现联合模型,利用BERT代替意图识别和槽位填充基线模型中的底层词向量表示模型。Qin等^[20]使用

预先训练的嵌入编码器以替换其注意力编码器,进一步提高了模型性能。

2 相关理论基础

2.1 Rasa NLU

NLU 处理消息是通过几个模块完成,这些模块连接执行,称为处理管道(Pipeline)。如图 1 所示,管道中的功能模块分为 4 种:分词器、特征提取器、分类器、实体识别模型。

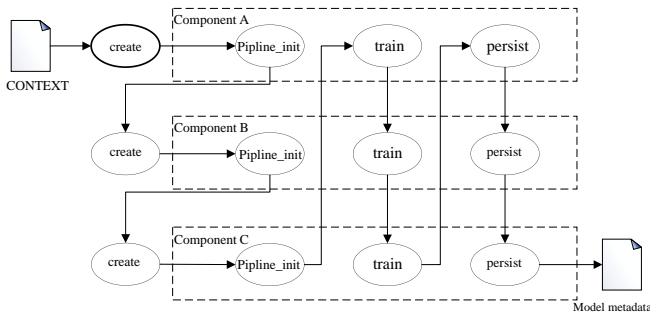


Fig. 1 Rasa NLU treatment pipeline

图 1 Rasa NLU 处理管道

2.2 RCNN

传统的特征表达方法有贝叶斯分类(NB)、支持向量机(SVM)、K近邻(KNN),但是这些方法忽略了上下文、语义和词序信息,而基于深度学习和词嵌入的方法在特征提取方面比传统的特征表示方法更有优势。在文本分类任务中,卷积神经网络(CNN)和循环神经网络(RNN)作为两种常用的深度学习方法的代表,被广泛应用于各种分类任务中,但是也有各自的局限性。CNN作为一种无偏模型,通过池化操作可以获得更重要的特征,但是在使用过程中卷积核大小难以固定,过小容易丢失数据,过大则空间参数过大。RNN作为一个有偏的模型,善于利用上下文关系,但是通常文本序列后面的单词有更大的权重,这并没有考虑到重要的单词可能出现在文本序列的任何地方。相比于前两种深度学习方法,RCNN的优势在于不需要像CNN一样需要通过窗口大小来设定对上下文的依赖长度,同时还可以较为均匀地利用单词的上下文信息,解决在RNN中后面的单词比前面的单词影响力更大的缺点^[21]。

递归卷积神经网络(RCNN)用于文本分类,其结构如图 2 所示。该模型采用双向循环结构,比传统的基于窗口的神经网络具有更少的干扰,可以最大限度地提取上下文特征。并且,RCNN通过池化层自动判断哪个特征具有更重要的作用。在RCNN中,词语的嵌入主要由 3 部分特征向量连接组成,3 部分特征向量分别是词左边的文本信息(lc)、词本身信息(e)以及词右边的文本信息。

$$x_i = [c_l(w_i); e(w_i); c_r(w_i)] \quad (1)$$

其中, w 代表一个词语, $e(w_i)$ 是通过处理后得到的向量, c_l 是融合当前词前面的词汇信息, c_r 是融合当前单词之

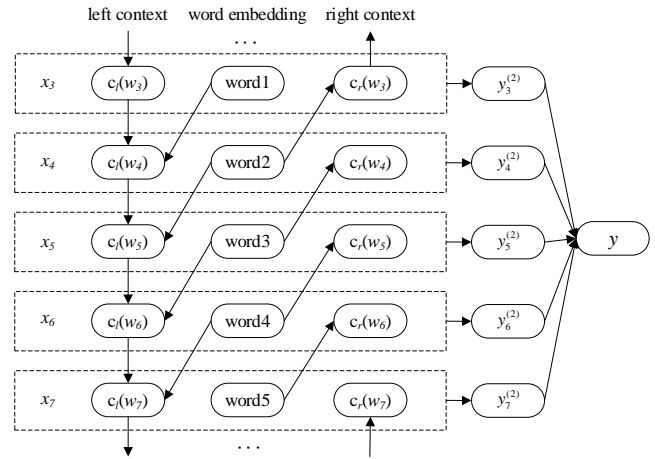


Fig. 2 Structure of recurrent convolutional neural network

图 2 递归卷积神经网络结构

后的信息。每一个 x_i 都用这种编码方式,融合了上下文信息并进行长期的相关预测。

$$c_l(w_i) = \int W^{(l)} c_l(w_{i-1}) + W^{(sl)} e(w_{i-1})] \quad (2)$$

$$c_r(w_i) = \int W^{(r)} c_l(w_{i+1}) + W^{(sl)} e(w_{i+1})] \quad (3)$$

其中, f 代表激活函数。非线性激活函数通过多层感知机处理并加入 \tanh 激活函数,可以得到 $y^{(2)}$,即每个单词类别的得分向量:

$$y_i^{(2)} = \tanh(W^{(2)} x_i + b^{(2)}) \quad (4)$$

使用最大池化得到 $y^{(3)}$:

$$y_i^{(3)} = \max y_i^{(2)} \quad (5)$$

经过上一步处理后再通过多层感知器和 softmax 层后得到最终得分向量:

$$y_i^{(4)} = W^{(4)} y_i^{(3)} + b^{(4)} \quad (6)$$

$y^{(4)}$ 经过 softmax 函数之后,将转化为概率:

$$P_i = \frac{\exp(y_i^{(4)})}{\sum_{k=1}^n \exp(y_k^{(4)})} \quad (7)$$

2.3 BIGRU-CRF

门控制循环单元和条件随机场(BIGRU-CRF)算法是一种实体识别模型,用于意图检测之后的槽位填充任务(其结构见图 3),本文槽位填充任务被转换为命名实体识别任务。BIGRU-CRF 方法首先通过预训练模型获得每个单词的向量表示,再由 BiGRU 作进一步的语义编码,最后输出到 CRF 层预测最大概率序列标签。CRF 可以通过标签之间的相邻关系获得全局最优的标签序列,对序列 $x = (x_1 + x_2 + \dots + x_n)$ 通过 BiGRU 进行标注输出序列标签 $y = (y_1 + y_2 + \dots + y_n)$,经过计算可以得出给定输入和输出之间的匹配分数。

$$s(s, y) = \sum_{i=1}^n (W_{y_{i-1}, y_i} + P_{i, y_i}) \quad (8)$$

P_{i, y_i} 是字符标签 y_i - th 的分数, $W_{i, j}$ 是标签的转换分数 (W 是转移矩阵)。

$$P_i = W_s h^{(i)} + b_s \quad (9)$$

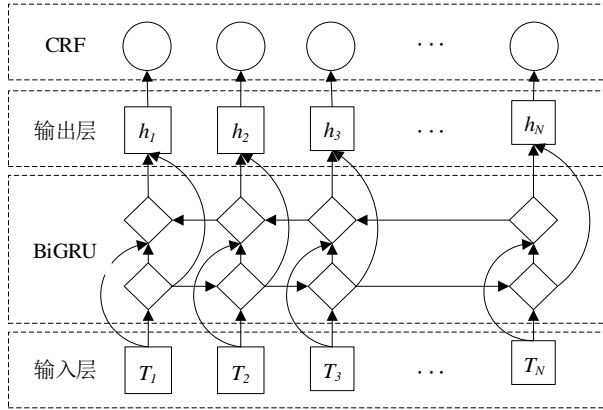


Fig. 3 BiGRU-CRF model structure

图3 BiGRU-CRF的结构

$h^{(t)}$ 是上一层 t 时间输入数据 x^t 的隐藏状态。参数是 CRF 的权重矩阵和最大条件似然估计,训练集是 x_i, y_i , 似然函数求解公式如下:

$$L = \sum_i^n \log(P(y_i|x_i)) + \frac{\lambda}{2} \|\theta\|^2 \quad (10)$$

P 代表输入序列对应于预测序列的概率:

$$p(y|x) = \frac{e^{s(x,y)}}{\sum_{y \in Y_t} e^{s(x,y)}} \quad (11)$$

因此,在进行意图识别之后可以通过 BiGRU+CRF 模型进行槽位填充。

2.4 基于混合技术的语义相似度算法

FinBERT 模型是国内首个针对金融领域大规模语料训练出的中文 BERT,该预训练模型是基于标准的 BERT 架构而建立,如图 4 所示。

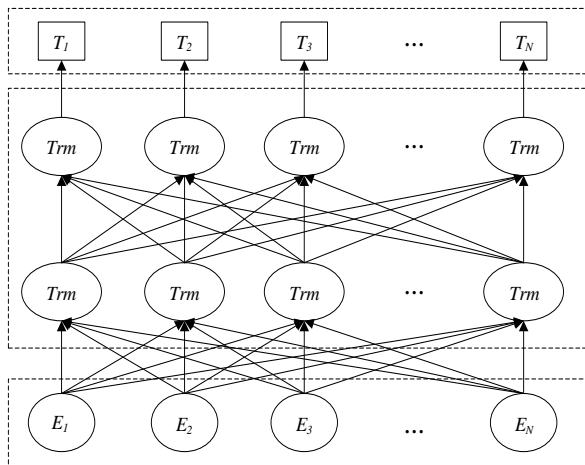


Fig. 4 BERT pretrained language model structure

图4 BERT预训练语言模型结构

近年来,预训练在 NLP 领域中非常流行,它构建了大量的预训练目标,以帮助模型更好地捕获语言知识和语义信息。此外,FinBERT 还通过多任务自监督预训练学习,不断更新预训练模型。与传统的预训练模型相比,可以同时一般在语料库和金融领域语料库上进行训练。在微调阶段,首先用预先训练好的参数初始化模型,然后对特定

于任务的监督数据进行微调。在多种金融数据集上进行实验,如在金融 SBD、SA 和 QA 数据集上,FinBERT 都取得了不错的结果,证明了该方法的有效性。

3 模块设计

3.1 数据预处理

目前,面向特定领域公开的中文数据集还很少,因此在进行 NLU 模块设计前,要先针对特定领域进行语料收集和文本预处理。之后结合智能催收电话系统的实际业务,对 NLU 模块进行具体设计,主要包括用户意图设计、槽位填充设计,再进行模型训练。语料收集流程如图 5 所示。

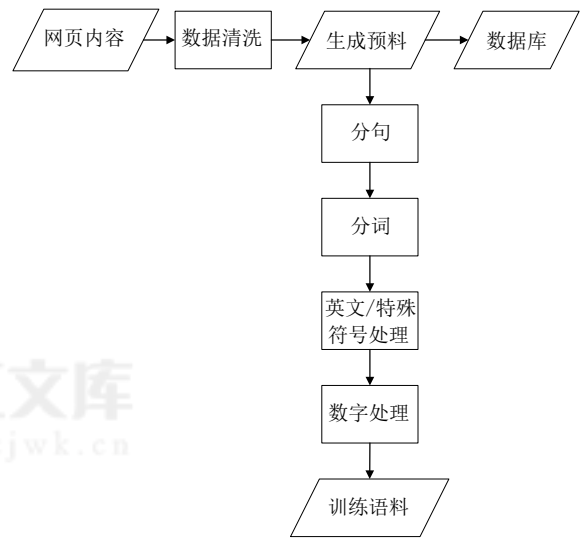


Fig. 5 Corpus preprocessing process

图5 语料预处理流程

在收集语料前,首先根据已有的数据和业务要求,定义必不可少的用户意图和系统动作,再从收集的语料库中抽取对应的数据进行训练。本文是针对金融催收领域主题的语料库,主要通过网络爬虫的方式爬取自己所需要的语料数据。本文采用 python 语言中的 Request 库模拟 HTTP 协议向新浪微博、中国金融新闻网、金融行业资讯网、每日金融新闻网等财经网站发送请求,获取相关的 HTML 页面。并且,还有一些金融公司对话日志,其中含有大量人工客服与客户的对话记录,这些对话可以作为智能催收系统中语言理解的训练语料。此外,可以从用户日志中选取一些催收机器人对应的场景,人工构造一部分用于项目冷启动,形成相应场景下的对话数据。通过上述途径,最终获得了 80 万份左右的原始语料。

收集到的文本形式多种多样,常见的有 TXT 文本、HTML 文本、XML 文本、Word 文本、Excel 文档等。HTML 文本中存在很多文本标签,例如“body”“div”“p”等。在这些标签的清洗过程中使用 Python 中的 BeautifulSoup 库和正则表达式库,完成对 HTML 页面的解析和清理工作。在对语料的标注过程中,本文主要是使用 Rasa 提供的语料标注

工具 Rasa NLU Trainer。标注好的语料总条数为 40 074 条,用户意图共有 28 类。数据集按照 9:1 的比例划分训练集和测试集进行 5 折交叉验证,选出最优超参数。

熵简科技于 2020 年提出基于 BERT 架构的金融领域预训练语言模型 FinBERT1.0,该模型在多个金融领域的下游任务中显著的性能提升,在没有任何调整的情况下 F1 值调高 2%~5.7%。本文尝试将 FinBERT 加入到意图识别与槽位填充任务中,并与原始 BERT 模型的结果进行比较。

3.2 意图识别模块设计

意图是对用户一句话表达的判断,判断用户想要完成什么任务^[22]。意图识别模块在设计过程中将模块设计为不可再分的类别,大部分的用户意图集中在“accept”“affirm”“autoDeduct”“delaypay”上,由此可以看出,在催收过程中,用户还是愿意还款或者有还款意向。意图“family”“refusepay”和“ask_bank_account”等,虽然在数据集中占比较少,但是保证了用户表达的多样性,有助于解决对话系统在突发场景下的切换问题。意图识别模块的主要流程是首先将文本输入到模型中,利用 FinBERT 进行预训练将输入文本转化为向量,在 FinBERT 训练的最后一层输出加上权重作为 RCNN 模型的输入,在 RCNN 模型中对文本的深层语义特征进行特征提取,并输出意图识别的结果。同时,将 FinBERT 模型得到的向量表示也作为槽位填充模型的输入。

3.3 槽位填充模块设计

在与用户的对话过程中,总有一些关键信息,需要深刻理解才能完成用户指令。槽位填充就是将用户表达的槽位内容识别为结构化的信息,例如时间、银行卡信息等。当用户想要知道自己账户逾期利息时,就需要知道很多用户的必要信息。在槽位设计过程中需遵守以下规则:①槽位需要被解释;②槽位是核心信息,有特定含义,需要在执行任务时使用;③槽位可以被清晰识别。在本文设计的智能催收系统中,表 1 列出了所有槽位填充情况。在本文提到的 FinBERT-BiGRU-CRF 模型中,整个模型分为 3 部分,首先使用 FinBERT 模型获得输入文本的向量表示,将得到的每个字的向量输入 BiGRU 中并对语义进行编码,最后输出到 CRF 层并输出最大概率的标签序列。

Table 1 Slot filling statistics

表 1 槽位填充统计

| 槽位 | 类型 | 解释 |
|-----------------|-------------------|-------------|
| story_end | Boolean slot | 故事是否结束的标志 |
| pay_status | Boolean slot | 支付状态:是否还款 |
| is_any_answer | Boolean slot | 是否启动任意回答的标志 |
| vf_story_end | Boolean slot | 验证身份的环节是否结束 |
| is_self | Boolean slot | 是不是用户本人 |
| num_doubt_robot | Unfeaturized slot | 用户质疑机器人的次数 |
| other | Boolean slot | 其他填槽 |

3.4 其他模块设计

在对话管理模块中一个意图对应着一个对话响应,系统要执行的动作主要根据 NLU 模块得到的用户意图以及相应的槽值进行设计。本文所使用的 Rasa Core 中允许多个意图对应一个或者多个系统动作以简化故事结构。在模块设计时,以“HS”开头的系统动作代表对应用户意图可以直接回复的动作如“HS_repeat”;以“QR”开头的动作代表系统需要主动对用户提问的动作;以“JS”开头的系统动作代表不同场景下对话的结束动作;以“Action”开头的系统动作代表需要进行槽位填充的自定义动作。

话术模板设计主要针对系统动作而定,用作系统对用户的具体响应输出。考虑到保护个人隐私和信息安全,对涉及到的相应人名、公司名、金额、电话、账户信息等,均用“XX”表示。如 QR 类模板:“请问您是本人吗?/请问这是 XX 的电话吗?”;HS 类模板:“如果你不想逾期记录恶化,麻烦尽快处理一下好吗?/那你等下把钱存到银行卡里,自己登录我们的 APP 还款,没问题吧!”;JS 类模板:“我们的通话是全程录音的,建议您登录 XX 借款 APP 选择充值还款的方式,在两小时还清款项,避免产生不良记录,感谢您的接听,再见。”

3.5 模型训练

通过对词向量的特征构造与选择,得到一个面向金融领域的预训练词向量模型,对训练语料进行相应标注和定义。训练过程中标注好的语料共有 40 074 条,用户意图共被划分为 28 类。将数据按照 9:1 的比例划分训练集和测试集,在训练集中按照 9:1 的比例划分出验证集,然后对训练集进行五折交叉验证。本文大多数实验都是在 Intel Xeon E5-2678 v3 CPU、双 Nvidia GeForce GTX 1080 Ti GPU 的环境下进行测试。首先加载预训练词向量模型,与 BERT 模型一样,FinBERT 模型也分为两种,即 FinBERT-large 和 FinBERT-base。两种模型网络结构大致相同,部分参数略有不同,这里使用的是 FinBERT-Base 模型,其中学习率是 $5e-5$,每批训练数据为 16,有 768 个隐藏神经元,编码器的层数以及自注意力头数均为 12。在 RCNN 模型中,词嵌入维度是 768,隐藏节点数是 256,学习率为 0.001,每批数据大小为 128,随机失活率为 0.1,epoch 设置为 100。随着训练次数的逐渐增多,准确率不断上升,损失率逐渐下降,最后在经过 100 次的 epoch 后,准确率在验证集和训练集开始震荡收敛。验证集准确率达 96% 左右,同时损失率也下降到 0.1,此时模型已达到最优。

4 自然语言理解模块测试与分析

在实验中除与传统模型对比外,还选取了 NLU 模块的准确率、召回率 F_1 值以完成 NLU 模块的测试与分析。选择准确率、召回率、 F_1 值对意图识别进行评价,选用 F_1 值对槽位填充模块进行评价。按照 9:1 的比例将数据划分为

训练集和测试集,遇到小数点按四舍五入的原则取整。对话系统各模块评价指标众多,在自然语言理解模块中,对意图识别和槽位填充识别进行精确率(P)、召回率(R)及 F_1 得分评测。其中, TP 是模型成功识别为正确的样本数, FP 是模型识别到的正确样本但实际是错误的样本个数, FN 是模型预测为错误样本但实际正确的样本个数。

$$P = \frac{TP}{TP + FP} * 100\% \tag{12}$$

$$R = \frac{TP}{TP + FN} * 100\% \tag{13}$$

$$F_1 = \frac{2 * P * R}{P + R} * 100\% \tag{14}$$

4.1 意图识别模块测试与评估

在意图识别任务中,为了验证本文 FinBERT 模型效果,加入与常见神经网络模型、BERT 模型的比较。为了验证 RCNN 的有效性,本文还加入了面向特定领域的预训练模型与不同神经网络算法结合的比较。不同模型在本文数据集上的比较结果如表 2 所示。

Table 2 Comparison of intention recognition results
表 2 意图识别实验结果比较

| 意图识别模型 | P (%) | R (%) | F ₁ (%) |
|--------------|-------|-------|--------------------|
| RNN | 83.48 | 83.28 | 83.38 |
| CNN | 82.54 | 82.16 | 82.35 |
| RCNN | 83.78 | 83.56 | 83.67 |
| BERT | 92.15 | 91.79 | 91.97 |
| FinBERT | 93.36 | 93.52 | 93.44 |
| FinBERT-RNN | 94.40 | 94.42 | 94.41 |
| FinBERT-CNN | 94.45 | 94.01 | 94.23 |
| FinBERT-RCNN | 95.81 | 95.69 | 95.75 |

在常见神经网络模型中的输入使用了 Word2vec 工具中的 Skip-gram 模型进行词向量表示。由表 2 可以看出,以 CNN 模型为基线,在与原始预训练语言模型 BERT 比较中,后者可以显著提高意图识别任务中的精确率、召回率、 F_1 值等。当面向特定领域的预训练语言模型与常见神经网络模型相结合后,实验效果进一步提升,RCNN 算法在与之前算法的比较中也展现出了自身优势,最终取得了 95.75% 的 F_1 值,高于基线 13.4%,验证了本文提出的算法模型性能更好。

4.2 槽位填充测试

表 3 是不同槽位填充模型在本文数据集上的实验结果比较。共享意图识别模块中预训练后的向量表示,本文通过 RasaNLU 内置的同义词抽取和正则表达式功能,利用 FinBERT+BiGRU+CRF 算法模型完成实体别任务。槽位填充模块的基线模型使用 Word2vec 工具中的 Skip-gram 模型对词向量进行表示,然后使用 BiGRU+CRF 方法完成槽位填充任务。在分别加入 BERT 模型及 FinBERT 模型后,实验的各项指标均有所提升。本文提出的 FinBERT+BiGRU+CRF 算法模型最终取得了 95.88% 的 F_1 值,

高于基线 7.31%。

Table 3 Comparison of experimental results of slot filling
表 3 槽位填充实验结果比较

| 槽位填充模型 | P (%) | R (%) | F ₁ (%) |
|-------------------|-------|-------|--------------------|
| BiGRU +CRF | 89.41 | 87.75 | 88.57 |
| BERT+BiGRU+CRF | 93.19 | 92.75 | 92.97 |
| FinBERT+BiGRU+CRF | 95.90 | 95.87 | 95.88 |

5 结语

本文基于开源机器学习框架 Rasa,改进了其中的 NLU 模块,采用最新提出的面向金融领域的预训练语言模型 FinBERT 与 RCNN 算法相结合的方式完成意图识别任务,共享意图识别任务中的预训练向量结合 BiGRU+CRF 的联合模型完成槽位填充任务。实验表明,本文算法的模型性能更优。后续将进一步丰富领域内的对话语料,引入闲聊功能以提升系统稳定性,并提高用户体验。

参考文献:

[1] YANG Y, CHEN L. Application of artificial intelligence in credit card collection[J]. China Credit Card, 2018(9): 9-13.
杨勇,陈玲. 人工智能在信用卡催收中的应用[J]. 中国信用卡, 2018(9): 9-13.

[2] HE Q. Development and application of artificial intelligence technology [J]. Electric Power Information and Communication Technology, 2017, 15(9): 32-37.
贺倩. 人工智能技术的发展与应用[J]. 电力信息与通信技术, 2017, 15(9): 32-37.

[3] CHE W X, ZHANG W N. Overview of man-machine dialogue system[J]. Artificial Intelligence View, 2018(1): 76-82.

[4] YE Y L, CAO B, FAN J, et al. Coarse-grained intent recognition method for task-oriented multi-turn dialogue[J]. Journal of Chinese Mini-Micro Computer Systems, 2018(9): 9-13.
叶银雷,曹斌,范菁,等. 面向任务型多轮对话的粗粒度意图识别方法[J]. 小型微型计算机系统, 2020, 41(8): 1620-1626.

[5] RAJPURKAR P, JIA R, LIANG P. Know what you don't know: unanswerable questions for SQuAD[C]// Meeting of the Association for Computational Linguistics, 2018: 784-789.

[6] RAJPURKAR P, ZHANG J, LOPYREV K, et al. Squad: 100 000+ questions for machine comprehension of text[C]// Austin: the 2016 Conference on Empirical Methods in Natural Language Processing, 2016.

[7] REDDY S, CHEN D Q, CHRISTOPHER D. CoQA: a conversational question answering challenge[J]. Transactions of the Association for Computational Linguistics, 2019 (7): 249-266.

[8] CUI L Y, WU Y, LIU S J, et al. MuTual: a dataset for multi-turn dialogue reasoning[C]// Proceedings of the 58th Conference of the Association for Computational Linguistics, 2020: 1406-1416.

- [9] HEMPHILL C T, GODFREY J J, DODDINGTON R, et al. The ATIS spoken language systems pilot corpus [C]// Proceedings of the darpa speech and natural language workshop, 1990: 96-101.
- [10] CHEN T T, LIN M, LI Y L. Joint intention detection and semantic slot filling based on BLSTM and attention [C]// Honolulu: 2019 IEEE 4th International Conference on Cloud Computing and Big Data Analysis (ICCCB-DA), 2019.
- [11] KIM Y. Convolutional neural networks for sentence classification [C]// Doha: the 2014 Conference on Empirical Methods in Natural Language Processing, 2014.
- [12] XU P Y, SARIKAYA R. Convolutional neural network based triangular CRF for joint intent detection and slot filling [C]// Olomouc: Automatic Speech Recognition proceedings of and Understanding (ASRU), 2013.
- [13] VU N T. Sequential convolutional neural networks for slot filling in spoken language understanding [C]// Paris: 17th Annual Conference of the International Speech Communication Association, 2016.
- [14] LIU Z H, WU G I, LIN Z J, et al. Attention-informed mixed-language training for zero-shot cross-lingual task-oriented dialogue systems [C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2020: 8433-8440.
- [15] ZHANG S Y, JING J J, HE C X, et al. A novel slot-gated model combined with a key verb context feature for task request understanding by service robots [J]. IEEE Access, 2019, 7: 5937-5947.
- [16] WANG Y, SHEN Y L, JIN H X. A Bi-model based RNN semantic frame parsing model for intent detection and slot filling [DB/OL]. <https://arxiv.org/abs/1812.10235>.
- [17] WU Q K, PENG D L. MTL-BERT: a multi-task learning model utilizing bert for Chinese text [J]. Journal of Chinese Mini-Micro Computer Systems, 2021, 42(2): 291-296.
- 武乾坤, 彭敦陆. MTL-BERT: 一种结合 BERT 的中文文本多任务学习模型 [J]. 小型微型计算机系统, 2021, 42(2): 291-296.
- [18] CHEN Q, ZHUO Z, WANG W. BERT for joint intent classification and slot filling [DB/OL]. <https://arxiv.org/abs/1902.10909>.
- [19] ZHAN B H. Research on joint algorithm of intention recognition and slot filling in task-based dialogue [D]. Beijing: Beijing University of Posts and Telecommunications, 2020.
- 战保行. 任务型对话中意图识别和槽填充的联合算法研究 [D]. 北京: 北京邮电大学, 2020.
- [20] QIN L B, CHE W X, YANG M L, et al. A stack-propagation framework with token-level intent detection for spoken language understanding [DB/OL]. <https://arxiv.org/abs/1909.02188>.
- [21] LI Y C, QIAN L F, MA J. Early detection of micro blog rumors based on BERT-RCNN [J]. Information Studies: Theory & Application, 2021, 44(7): 173-177, 151.
- 李悦晨, 钱玲飞, 马静. 基于 BERT-RCNN 模型的微博谣言早期检测研究 [J]. 情报理论与实践, 2021, 44(7): 173-177, 151.
- [22] HU G M, JIANG L. Gated slot model integrating intention list query mechanism [J]. Software Guide, 2021, 20(9): 51-55.
- 胡光敏, 姜黎. 融合意图列表查询机制的门控槽模型 [J]. 软件导刊, 2021, 20(9): 51-55.

(责任编辑: 孙 娟)