聊天机器人中用户出行消费意图识别方法

钱岳, 丁效, 刘挺*, 陈毅恒

哈尔滨工业大学计算机科学与技术学院信息检索研究中心, 哈尔滨 150001

* 通信作者. E-mail: tliu@ir.hit.edu.cn

收稿日期: 2016-12-31; 接受日期: 2017-03-06; 网络出版日期: 2017-06-09

国家重点基础研究发展计划 (973) (批准号: 2014CB340503) 和国家自然科学基金 (批准号: 61472107, 61632011) 资助项目

摘要 聊天机器人中的出行消费意图是指用户为了满足出行的需要, 通过文本表达出对出行类产品或者服务的购买意愿. 识别出用户的消费意图可以进行相应的产品推荐, 增强用户体验. 传统的消费意图识别主要使用基于模板匹配或者基于人工特征集合的机器学习方法, 这类方法费时费力, 扩展性不强. 本文将出行消费意图识别任务看成一个分类问题, 结合深度学习方法识别用户的出行消费意图, 该方法不需要人工构造特征集合或匹配模板. 具体而言, 本文构建了基于卷积的长短期记忆神经网络 (Convolutional-LSTM) 模型进行出行消费意图识别, 首先通过卷积神经网络 (CNN) 对用户的聊天文本进行特征抽取, 随后进行特征组合并送入长短记忆神经网络 (LSTM) 进行特征表示学习, 最后输出分类结果. 实验结果表明, 在出行消费意图识别任务上, 基于 Convolutional-LSTM 的模型在 F 值上优于最好的基线方法 2 个百分点.

关键词 出行消费意图识别, 机器学习, 深度学习, 卷积神经网络, 长短期记忆网络

1 引言

随着社会的日益信息化, 人们希望能用自然语言与计算机交流, 于是聊天机器人应运而生. 在和机器人聊天的过程中, 用户产生了大量的数据. 这些数据包含了用户丰富的信息, 展示了用户的需求和意图, 识别出用户的出行意图可以进行相应产品的推荐, 使得聊天机器人变得更加智能, 增强用户体验. 基于此, 本文提出了聊天机器人中用户出行消费意图的识别工作. 所谓出行消费意图是指用户为了满足出行的需要, 通过聊天文本表达出的对出行类产品或服务的购买意愿. 准确地识别出用户的出行消费意图, 可以了解用户的出行意愿, 有针对性地进行消费产品的推荐, 更好地服务于用户.

出行消费意图识别工作面临着对聊天文本语义理解不够深入, 无法在用户较短的对话文本中准确理解用户意图等诸多挑战^[1]. 以往的消费意图识别工作主要可以分为基于模板匹配的方法和基于有指导的分类方法^[2,3]. 前者主要通过挖掘句子中是否出现意图模板来确定句子最终是否具有消费意图.

引用格式: 钱岳, 丁效, 刘挺, 等. 聊天机器人中用户出行消费意图识别方法. 中国科学: 信息科学, 2017, 47: 997–1007, doi: 10.1360/N112016-00306
Qian Y, Ding X, Liu T, et al. Identification method of user's travel consumption intention in chatting robot (in Chinese). Sci Sin Inform, 2017, 47: 997–1007, doi: 10.1360/N112016-00306

后者则将 n-gram、模板等作为分类的特征, 联合其他特征训练分类器来完成句子的消费意图识别. 然而, 无论是基于模板匹配还是有指导的方法, 在应用中都有一定的局限性. 首先, 研究者们通常认为出现模板的句子即具有消费意图, 故而人工定义或自动抽取高频短语作为模板使用. 但在实际应用中, 基于模板的方法覆盖率不高, 效果不理想. 其次, 由于用户聊天文本的特殊性, 其文本长度一般很短, 通过构建特征的方式不能很好地学习到文本的深层语义信息, 这也就导致了不能准确地理解用户的消费意图.

本文把用户的出行消费意图识别任务看成一个分类问题, 首先尝试了基于支持向量机 SVM 的传统机器学习方法. 考虑到领域移植的便捷性和人工构造特征的诸多不便, 所以采用端到端的深度学习方法. 首先尝试基于卷积神经网络 CNN 的方法, 该方法自动从文本中抽取隐含的特征后进行出行消费意图的识别, 但是 CNN 模型对文本时序关系的刻画能力不强. 因此, 为了能更好地利用用户聊天文本的上下文时序语义信息, 进而使用 LSTM 学习句子语义表示后进行消费意图的识别, 但 LSTM 比较依赖外部的原始词向量输入. 基于上述情况, 根据任务和用户聊天文本数据的特殊性, 结合了 CNN 和 LSTM 的优点, 构建了一种基于卷积的长短期记忆网络 Convolutional-LSTM, 使用 CNN 从文本中抽取特征, 并经过特征组合后使用 LSTM 学习聊天文本的最终表示. 该网络既考虑了句子的隐含特征信息, 又考虑了句子中词语之间的时序语义关系.

实验结果表明, 在出行消费意图识别任务上, 本文采用的混合模型在各项指标上均优于传统的机器学习方法, 以及基于 CNN 和基于 LSTM 的深度学习方法至少 2 个百分点, 能够有效识别用户的出行消费意图.

2 相关研究

出行消费意图识别旨在判断一个用户发布的文本或表现的行为是否具有某种出行的消费意图, 如订机票、火车票和酒店等. 由于聊天机器人发展的时间较短, 相关的技术尚不成熟, 因而基于用户聊天文本的出行消费意图识别工作还比较少. 虽然消费意图识别相关研究工作刚刚兴起, 但是正在渐渐地吸引学术界和工业界的广泛关注^[4]. 消费意图识别不仅有助于了解用户的潜在购买需求, 还能辅助事件的预测, 判断事件的走向^[5]. 该项工作可以认为是计算机科学、心理学和市场营销学的多学科交叉领域.

Dai 等^[6]提出了识别在线商业意图 (online commercial intention, OCI) 研究任务. 该方法认为, 如果一个查询点击了具有商业意图的网页, 则此查询也具有商业意图. 具体地, 他们将查询商业意图检测定义为二元分类问题, 首先基于网页中的关键词、标签等特征设计分类器, 识别具有商业意图的网页之后将查询映射到二值函数上 (即是否有商业意图). Ashkan 等^[7]利用用户的点击流信息、查询自身特征以及搜索结果页面的内容特征检测用户的查询商业意图. 实验结果表明, 将这 3 类特征组合起来可以有效地检测出用户的查询商业意图. 此外, Guo 等^[8]结合用户的鼠标点击和滚动等用户搜索行为建模, 对两种搜索意图“寻找” (research) 和“购买” (purchase) 进行分类. Kooti 等^[9]通过对邮件信息进行分析, 发现消费行为受到年龄、性别等因素的影响, 并基于此进行用户意图以及消费行为的预测. Goldberg 等^[10]提出了“buy wish”的概念. 他们从用户在互联网上发布的愿望清单和产品评论中的内容来识别其中的消费意图. Chen 等^[11]在研究中由于消费意图语料的匮乏, 他们在消费意图表达具有相似性的假设下提出了跨领域迁移学习 (transfer learning) 的消费意图检测方法, 该方法对于解决语料匮乏的问题有一定帮助.

近些年随着社会媒体技术的发展和普及, Twitter 和新浪微博等社会媒体成为最普遍的信息发布、

表 1 用户常见表达举例
Table 1 Examples of common user expression

No.	Intentional dialogue	Unintentional dialogue
1	下周出差飞哈尔滨	你叫什么名字
2	附近有住宿的地方吗	今天的天气如何
3	到北京的飞机	我刚下飞机, 到北京了
4	火车硬座票还有吗	坐了一天的火车, 还是硬座
5	和颐酒店大床房	和颐酒店事件

传播和共享的工具之一. 在这些用户生成的数据中蕴含了丰富的消费需求信息. Liu 等^[3]提出了基于弱监督的图排序算法, 该方法适用于数据总量较大, 已标注数据较少的情况. Ding 等^[12]提出了基于领域自适应的消费意图检测方法. 具体而言, 给定一条微博文本, 该方法首先基于领域自适应的卷积神经网络模型做二元分类, 判断其是否含有用户消费意图; 随后, 对于有消费意图的微博文本再抽取其中的需求词; 最后, 根据该需求词去事先构造好的消费意图事理图谱中检索可以满足消费意图的产品推荐给用户. Duan 等^[13]提出了基于社区问答的消费意图识别和挖掘方法, 通过挖掘问题与回答之间的需求对应关系, 把消费对象的挖掘问题转化为搭配问题. Wang 等^[14]将 Twitter 中用户潜在的意图划分为 6 个类别, 并通过半监督学习的方法进行消费意图识别. Yang 等^[15]提出了人们日常生活中经常出现的 12 种普遍需求, 通过模板匹配方法获得了购买产品的用户, 然后基于 Twitter 中的 unigram 特征、WordNet 中词的语义特征和表达需求的动词特征来训练分类器完成对用户消费意图的识别.

区别于传统的消费意图识别方法, 本文尝试使用端到端的模型进行意图识别工作. 并且根据聊天文本的特性, 结合用户的表达方式和特点, 通过对不同的深度学习模型进行整合, 进行聊天机器人中用户的消费意图识别.

3 问题描述

本文主要研究 3 个领域的出行意图识别任务, 即订机票意图、订火车票意图和订酒店意图, 常见的聊天文本如表 1 所示. 进行用户出行意图识别的基本目标是根据用户在聊天中表达的语义信息来确定用户聊天文本的出行消费意愿, 即判断用户在文本中表达了哪一类出行意图.

对于表 1 中前两组例子, 直观上看, 有出行意图和无出行意图的文本表达在用词、句式上存在较大的差异, 可以直接使用一些模板匹配的方式进行识别. 但是在聊天语料处理过程中发现, 由于聊天系统对话的特殊性, 用户的表达一般较短, 一些无意图的用户聊天文本在用词、句式上和有意图文本比较接近, 通过一些模板或者关键词特征的方法不容易区分. 从表 1 中的第 3~5 组例子可以看出, 每一组的两句对话在用词或者句式上都具有很高的相似性, 但却属于两类不同的意图领域. 对于第 3 组例子, 尽管二者都具有相同的“动词 + 地点”的动宾关系句法结构, 并且都包含了“飞机”这一关键词, 但是有意图文本希望查询到某地的航班信息, 而无意图文本则是表达了用户刚刚下飞机; 对于第 4 组例子, 有意图对话表达想去北京要订火车票, 而无意图表达只是表示自己坐了一天的火车; 对于第 5 组例子, 有意图文本很短, 只有 3 个词语, 但是却包含了“酒店”、“大床房”等关键词, 无意图文本虽然也包含了关键词, 但是表达的却是一个新闻事件, 并不是希望订酒店.

由此可见, 在聊天机器人中, 用户出行意图的识别具有很大的挑战, 如何从较短的, 而且是表达相近的用户聊天文本中准确地识别出用户的出行消费意图, 是一项十分具有挑战性的工作. 因而本文的

图 1 基于 SVM 的出行消费意图识别流程图

Figure 1 The flow chart of consumer intention recognition based on SVM

主要目标是识别出用户的出行消费意图, 准确地将这些相似的无意图文本进行区分, 进而理解用户表达的真实意图, 服务于后续的推荐工作.

4 模型

4.1 基于 SVM 的出行消费意图识别

以往的消费意图识别工作主要基于支持向量机 (SVM) 等传统机器学习方法^[16], 因此本文为了与前人方法进行更好的对比, 将基于 SVM 的出行消费意图识别方法作为基线方法. 在该实验中使用词特征作为模型输入, 首先对数据集进行分词等预处理后生成词表, 然后对数据集中的每条文本根据词表进行 one-hot 表示作为最终的模型输入. 该方法对出行消费意图的识别流程如图 1 所示.

首先对标注语料进行分词等预处理, 统计语料中不重复的词语生成词表并对数据进行向量空间表示, 进而将数据划分为训练集、开发集和测试集. 使用训练集进行模型的训练, 并在开发集上进行参数调整, 得到效果最优的模型后, 在测试集上评估模型的性能.

4.2 基于 CNN 的出行消费意图识别

CNN 模型已经证实文本分类等任务中具有良好的表现^[17]. 根据本文的问题描述, 出行消费意图识别本质上是一个四元分类问题, 因此尝试使用基于 CNN 的出行消费意图识别方法. 该模型如图 2 所示, 由输入层、卷积层、池化层和最终的 Softmax 分类层组成.

(1) **输入层.** 对于给定长度为 n 的用户聊天文本, 使用 $x_i \in \mathbb{R}^k$ 代表句子中第 i 个词维的词向量表示, 整个句子可以表示为 $S \in \mathbb{R}^{nk}$, 由句子中词向量进行连接而成, 如式 (1) 所示:

$$x_{1:n} = x_1 \oplus x_2 \oplus \cdots \oplus x_n, \quad (1)$$

其中, $\oplus$ 代表将词向量依次进行连接操作.

(2) **卷积层.** 卷积操作可以看作是基于滑动窗口的特征提取, 主要用来捕捉词语间的局部信息, 本文定义卷积核的长度为 h , 大小为 $w \in \mathbb{R}^{hk}$, 用于聊天文本中第 i 个词到第 $i+h-1$ 个词. 卷积核通过式 (2) 所示的卷积操作得到一个特征表示:

$$f_i = f(w_i \cdot x_{i:i+h-1} + b). \quad (2)$$

图 2 (网络版彩图) 基于 CNN 的出行消费意图识别流程图

Figure 2 (Color online) The flow chart of consumer intention recognition based on CNN

该卷积核依次对句子中所有的窗口 $\{x_{1:h}, x_{2:h+1}, x_{3:h+2}, \dots, x_{n-h+1:n}\}$ 进行滑动卷积, 得到一个 feature map $F \in \mathbb{R}^{n-h+1}$, 如式 (3) 所示:

$$F = [f_1, f_2, \dots, f_{n-h+1}]. \quad (3)$$

(3) 池化层. 本文采用 max-over-time pooling 进行池化, 取每个 feature map 中值最大的一维作为抽取出的最终特征. 该层主要用来将局部的上下文特征向量映射成固定长度的全局特征向量, 使得仅保留具备特征向量中最具有代表性的特征.

(4) Softmax 层. 本任务一共有 4 个输出, 即该用户文本是否具有订机票、火车票、酒店意图或者是无意图. 因此该模型 Softmax 层输出维度为 4. Softmax 之所以适用于本问题场景, 在于它的输出可以解释成条件概率, 根据条件概率的值来识别聊天文本到底属于哪一类出行意图.

4.3 基于 LSTM 的出行消费意图识别

基于 CNN 的出行消费意图识别虽然对文本进行深层特征抽取得到了文本最终表示, 但在特征抽取过程中没有很好地考虑文本的语序问题. 对于聊天机器人的用户对话文本, 由于其表达较短, 且词语的先后语序关系对文本的实际含义影响较大, 为了能更好地对文本的时序关系建模, 采用基于 LSTM 的出行消费意图识别方法. LSTM 模型由 Hochreiter 等^[18] 提出, 它不仅对文本的时序关系具有良好的建模能力, 而且当用户输入的文本较长时, 可以有效处理文本的长期依赖问题. 本文中的 LSTM 结构如图 3 所示, 该模型由输入层、长短期记忆层和 Softmax 分类层组成, 其中 Softmax 层与 CNN 模型中相同, LSTM 层由一系列重复的 Cell 组成, 得到最后一个 Cell 的输出送入 Softmax 层.

(1) 输入层. 该模型的输入层与 CNN 中不同的是, 词向量是按照顺序依次送入到对应的 Cell 中, 以保证聊天文本的时序性.

图 3 (网络版彩图) 基于 LSTM 的出行消费意图识别流程图

Figure 3 (Color online) The flow chart of consumer intention recognition based on LSTM

(2) 长短期记忆层. 长短期记忆层由一系列重复的 Cell 组成, 每个 Cell 接收上一时刻隐层 h_{t-1} 的输出以及当前时刻的输入 h_t . 每一个 Cell 由输入门 i_t 、遗忘门 f_t 和输出门 o_t 组成, 对于 LSTM 中的每一个神经元, 其工作流程如下:

输入门用来控制当前输入 x_t 以及前一时刻隐层的输出 h_{t-1} 进入新的 Cell 的信息量, 如式 (4) 所示, 其中 σ 代表 sigmoid 非线性函数,

$$i_t = \sigma(w_t \cdot x_t + w_t \cdot h_{t-1} + b_t). \quad (4)$$

遗忘门用来控制清除或者保留现有的信息量的程度, 其式为

$$f_t = \sigma(w_t \cdot x_t + w_t \cdot h_{t-1} + b_t). \quad (5)$$

Cell 变换更新: 变换输入和前一状态到最新状态, 其中 $\tanh$ 代表双曲线正切函数,

$$q_t = \tanh(w_q \cdot x_t + w_q \cdot h_{t-1} + b_q). \quad (6)$$

输出门用来控制输出的信息量, 如式 (7) 所示:

$$o_t = \sigma(w_o \cdot x_t + w_o \cdot h_{t-1} + b_o). \quad (7)$$

计算下一个时间戳的状态, 如式 (8) 所示, 其中 $\odot$ 代表数组元素依次相乘运算,

$$c_t = f_t \odot c_{t-1} + i_t \odot q_t. \quad (8)$$

最终的隐层输出: 使用一个对当前状态的 $\tanh$ 变换进行重变换,

$$h_t = o_t \odot \tanh(c_t). \quad (9)$$

4.4 基于 Convolutional-LSTM 的出行消费意图识别

CNN 可以从文本中获取深层的特征信息, LSTM 则可以对文本的时序关系建模. 为了充分利用二者的优点, 本文将 CNN 与 LSTM 模型进行整合, 使用基于 Convolutional-LSTM 的出行消费意图

图 4 (网络版彩图) 基于 Convolutional-LSTM 的出行消费意图识别模型

Figure 4 (Color online) The flow chart of consumer intention recognition based on Convolutional-LSTM

识别模型, 该模型于 2015 年由 Sainath 等^[19]提出, 并在语音、图像领域有了一些应用^[20~23], 结构如图 4 所示, 包含了输入层、卷积层、特征组合层、长短期记忆层以及最终的 Softmax 层. 本文将该模型应用在文本上, 其卷积过程与前人工作不同的是, 卷积核的宽度和词向量维度相同. 本任务中该模型的输入层、卷积层、Softmax 层结构和 CNN 中对应层结构相同, 长短期记忆层与 LSTM 中对应层结构相同. 特征组合层的每个向量通过将卷积后的每个 feature map 的第 i 维对应的元素顺序相连得到, 如式 (10) 所示:

$$C_i = F_1^i \oplus F_2^i \oplus \dots \oplus F_{n-h+1}^i. \quad (10)$$

得到 $n - h + 1$ 个组合向量, 这样每个向量根据卷积的先后顺序进行重新组合, 保证了文本的时序性. 进而将向量依次送入 LSTM 的 Cell 中. 得到最后一个隐层的输出作为最终的句子表示, 交由 Softmax 层得到识别结果.

4.5 参数设置

在超参数的设置上, CNN 实验中, 设置卷积核长度为 3, 个数为 100, Padding 长度为 20 个词; LSTM 实验中, 设置输出维度为 100; Convolutional-LSTM 模型中的参数与上述二者相同. 设置训练过程中使用 $\text{batch_size} = 32$. 为了防止数据过拟合, 使用 dropout 和 L2 正则化进行约束, 其中 dropout rate = 0.25, 应用于输入层与卷积层之间以及 LSTM 层与 Softmax 分类层之间. L2 正则化应用于最终的 Softmax 分类层.

表 2 意图识别实验结果
Table 2 Experimental results of intention recognition

Model	Precision	Recall	F-Measure
SVM	0.9116	0.7769	0.8153
CNN	0.9469	0.9139	0.9278
LSTM	0.9475	0.9106	0.9258
Convolutional-LSTM	0.9514	0.9442	0.9473

5 实验

5.1 实验数据

本文的实验数据来源于科大讯飞灵犀语音助手 2015 年 4 月的用户日志, 共计 13677844 条用户数据. 灵犀语音助手是一款智能语音软件, 提供了查询及购买机票、火车票和酒店等功能, 其用户表达习惯和聊天机器人中用户的表达相似. 本文使用一些关键词及模板规则对数据进行了预处理, 过滤掉部分无意图数据, 在预处理后的数据上由两名标注人员进行标注, 共标注 28795 条数据. 通过计算两名标注人员标注结果的 Pearson 相关系数 (Pearson correlation coefficient) 来验证标注的一致性. 依照上述的流程, 得到最终 Pearson 相关系数值为 0.88, 说明不同用户对出行消费意图的理解较为一致. 同时, 为了避免实验结果的偶然性, 采用 5 折交叉验证的方法, 即将数据等分为 5 份, 依次取其中的 1 份作为测试集, 其余 4 份作为训练集, 训练模型并进行测试.

5.2 评价指标

对于每个类别均采用 Precision (准确率)、Recall (召回率) 和 F-measure 来评价分类器的性能, 如式 (11) ~ (13) 所示:

$$\text{Precision} = \frac{TP}{TP + FN}, \tag{11}$$

$$\text{Recall} = \frac{TP}{TP + TN}, \tag{12}$$

$$\text{F-measure} = \frac{2 \times P \times R}{P + R}. \tag{13}$$

其中, 属于类 C 的样本被正确分类到类 C , 记这一类样本数为 TP; 不属于类 C 的样本被错误分类到类 C , 记这一类样本数为 FN; 属于类别 C 的样本被错误分类到类 C 的其他类, 记这一类样本数为 TN; 不属于类别 C 的样本被正确分类到了类别 C 的其他类, 记这一类样本数为 FP.

对于本任务的多分类问题, 首先计算每个类别的评价指标, 进而以宏平均值 (macro-average) 作为最终的评价指标.

5.3 实验结果与分析

为了避免实验结果的偶然性, 本文采用 5 折交叉验证的方式, 在不同模型下的实验结果如表 2 所示.

实验结果表明, 基于 SVM 的传统机器学习方法在识别的准确率、召回率和 F 值上效果较差. 采用深度模型 CNN 和 LSTM 后, 效果得到了显著提高, 尤其是当结合二者优点使用 Convolutional-LSTM 模型时, 其在测试集上的 F 值达到了 0.9473.

在出行消费意图识别任务上,传统的基于 SVM 的机器学习方法采用模板匹配或者基于词和句法结构的特征进行模型的训练,该方法不能完全表达原有的文本信息,仅仅通过考量词语或者特征的出现情况,无法刻画聊天文本中词语之间的时序关系和语义信息,对于用词或者句法结构相近的文本区分性较差,识别精度不高.并且由于模板和特征抽取的局限性,不能很好地覆盖用户多样性的意图表达,导致实验结果召回率较低.

CNN 通过使用卷积核对文本进行局部感知得到多个特征向量,并通过池化操作对卷积后的向量进行特征选择,对重要的特征信息进行保留,可以从句子中抽取出深层的特征信息进行识别,可以看作是对原文本进行潜在信息的转换,在需要挖掘文本深层信息的任务上具有良好的表现,这是因为使用 CNN 对聊天文本进行特征抽取并进行选择后再进行分类工作,可以充分考虑聊天文本中重要的隐含类别信息.

同时,本文考虑了用户聊天文本的上下文时序关系对语义的影响,采用了基于 LSTM 的深度学习方法, LSTM 通过循环的网络结构,并通过 Cell 的门控机制进行信息的保留和传递,可以将文本的上下文信息进行有效的刻画,该方法不仅对于句子的语义理解有一定优势,而且当用户输入的文本较长时,可以有效处理文本的长期依赖问题,因此在需要考虑上下文关系任务上具有良好的表现.本文使用 LSTM 首先进行聊天文本的表示学习,得到最后的隐层输出,进行出行消费意图分类,可以充分考虑文本的时序关系以及语义信息.

但上述二者都分别具有一定的局限性,首先 CNN 由于需要对卷积后的特征向量进行池化操作以保留重要的特征信息用于分类,在池化操作中需要有选择性地进行信息保留,因而该操作会丢弃卷积后的文本上下文信息,不能完全考虑聊天文本中词语之间的时序关系,因此对于文本的语义理解不够;而 LSTM 则需要依赖原始的词向量输入,对于聊天机器人中的出行消费意图识别这一新的任务,训练词向量的大规模语料较难获取,因而无法得到十分适合本任务的初始化向量表示,进而缺乏任务相关的深层次语义信息,因此该模型对聊天文本的刻画不充分.

所以本文将二者进行了结合,采用 Convolutional-LSTM 的模型首先通过卷积操作,对用户聊天文本进行局部特征信息抽取以获取深层文本信息,然后通过对不同卷积核卷积后生成的特征向量进行维度组合,进而送入 LSTM 进行表示学习.由于该模型省去了卷积后的池化操作,特征向量保留了文本卷积后的上下文关系,通过对不同特征向量的相同位置元素进行组合,生成新的向量集合代替原有的词向量输入,利用 LSTM 的循环输入结构以及上下文信息保留和传递机制,得到最终的文本向量表示.分别结合了 CNN 在特征抽取和 LSTM 在语义理解方面的优势.实验结果表明,Convolutional-LSTM 方法在消费意图识别任务上具有良好的表现.

6 结论

在聊天机器人中,识别出用户的出行消费意图对于理解用户潜在行为和产品推荐具有重要的意义.消费意图识别是近年来新兴起的研究方向,传统的消费意图识别主要使用模板规则匹配或者基于特征的机器学习方法,这类方法费时费力、扩展性差,而且效果往往不理想.本文将出行消费意图识别任务看作一个分类任务,首先尝试了基于 SVM 的传统机器学习方法,同时使用深度学习 CNN 以及 LSTM 模型识别用户的出行消费意图,并结合了二者的优点,构建了一种基于卷积的长短期记忆网络 Convolutional-LSTM,通过对卷积后的文本进行特征组合,依次送入 LSTM 中学习句子表示后进行出行消费意图识别.实验证明,在本任务上,Convolutional-LSTM 模型较基于 SVM 的传统机器学习识别方法有较大提升,并在 F 值上优于 CNN 和 LSTM 的深度学习方法 2 个百分点,对于聊天机器人中用户的出行消费意图识别是行之有效的.

参考文献

- 1 Fu B, Liu T. Consumption intent recognition for social media: task, challenge and opportunity. *Intell Comput Appl*, 2015, 5: 1–4 [付博, 刘挺. 面向社会媒体的消费意图识别: 任务、挑战与机遇. *智能计算机与应用*, 2015, 5: 1–4]
- 2 Fu B, Liu T. Weakly-supervised consumption intent detection in microblogs. *J Comput Inf Syst*, 2013, 9: 2423–2431
- 3 Liu T, Fu B, Chen Y H. Detecting consumption intention based on graph ranking in social media. *Sci Sin Inform*, 2015, 45: 1523–1535 [刘挺, 付博, 陈毅恒. 基于图排序的社会媒体用户的消费意图检测. *中国科学: 信息科学*, 2015, 45: 1523–1535]
- 4 Zhang F, Yuan N J, Lian D, et al. Mining novelty-seeking trait across heterogeneous domains. In: *Proceedings of the 23rd International Conference on World Wide Web*, Seoul, 2014. 373–384
- 5 Liu T, Ding X, Chen Y, et al. Predicting movie Box-office revenues by exploiting large-scale social media content. *Multimedia Tools Appl*, 2016, 75: 1509–1528
- 6 Dai H K, Zhao L, Nie Z, et al. Detecting online commercial intention (OCI). In: *Proceedings of the 15th International Conference on World Wide Web*, Edinburgh, 2006. 829–837
- 7 Ashkan A, Clarke C L A. Term-based commercial intent analysis. In: *Proceedings of International ACM SIGIR Conference on Research and Development in Information Retrieval*, Boston, 2009. 800–801
- 8 Guo Q, Agichtein E. Ready to buy or just browsing?: detecting web searcher goals from interaction data. In: *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, Geneva, 2010. 130–137
- 9 Kooti F, Lerman K, Aiello L M, et al. Portrait of an online shopper: understanding and predicting consumer behavior. In: *Proceedings of the 9th ACM International Conference on Web Search and Data Mining*, Insead, 2016. 205–214
- 10 Goldberg A B, Fillmore N, Andrzejewski D, et al. May all your wishes come true: a study of wishes and how to recognize them. In: *Proceedings of Human Language Technologies: the 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, Boulder, 2009. 263–271
- 11 Chen Z, Liu B, Hsu M, et al. Identifying intention posts in discussion forums. In: *Proceedings of Conference of the North American Chapter of the Association for Computational Linguistics—Human Language Technologies*, Atlanta, 2013. 1041–1050
- 12 Ding X, Liu T, Duan J, et al. Mining user consumption intention from social media using domain adaptive convolutional neural network. In: *Proceedings of Association for the Advancement of Artificial Intelligence*, Austin, 2015. 2389–2395
- 13 Duan J W, Chen Y H, Liu T, et al. Mining intention-related products on online Q&A community. *J Comput Sci Tech*, 2015, 30: 1054–1062
- 14 Wang J, Cong G, Zhao W X, et al. Mining user intents in Twitter: a semi-supervised approach to inferring intent categories for Tweets. In: *Proceedings of Association for the Advancement of Artificial Intelligence*, Austin, 2015. 318–324
- 15 Yang H, Li Y. Identifying user needs from social media. *IBM Res Div*, 2013. 1–10
- 16 Huang D, Huang Z H. Consumption pattern recognition system based on SVM. In: *Proceedings of the 4th International Conference on Intelligent Computation Technology and Automation*, Breckenridge, 2011. 79–82
- 17 Kim Y. Convolutional neural networks for sentence classification. *arXiv*: 1408.5882, 2014
- 18 Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput*, 1997, 9: 1735–1780
- 19 Sainath T N, Vinyals O, Senior A, et al. Convolutional, long short-term memory, fully connected deep neural networks. In: *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, London, 2015. 4580–4584
- 20 Shi X J, Chen Z R, Wang H, et al. Convolutional LSTM network: a machine learning approach for precipitation nowcasting. In: *Proceedings of Advances in Neural Information Processing Systems*, Montréal, 2015. 802–810
- 21 Wu L, Shen C, Hengel A. Convolutional LSTM networks for video-based person re-identification. *arXiv*: 1606.01609, 2016
- 22 Zhou C, Sun C, Liu Z, et al. AC-LSTM neural network for text classification. *Comput Sci*, 2015, 1: 39–44
- 23 Donahue J, Hendricks L A, Guadarrama S, et al. Long-term recurrent convolutional networks for visual recognition and description. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, London, 2015. 2625–2634

Identification method of user's travel consumption intention in chatting robot

Yue QIAN, Xiao DING, Ting LIU* & Yiheng CHEN

Research Center for Social Computing and Information Retrieval, School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China

* Corresponding author. E-mail: tliu@ir.hit.edu.cn

Abstract Travel consumption intention in chatting robot is the users in order to meet their travel needs, express the willingness to purchase a product or service. Identifying the user's intent to consume the product can be recommended to enhance the user's experience. Traditional consumer intention recognition methods are mainly based on template matching or artificial feature sets, which are time consuming, laborious, and hard to extend. In this paper, we regard the travel consumption intention recognition task as a classification problem and combine the deep learning method to identify the intention. The proposed method does not need to construct the feature set or match templates manually. Specifically, this study uses the convolutional long short-term memory neural network (LSTM) model to identify the travel consumption intention. First, the feature extraction is carried out by creating a convolution neural network (CNN) of the user's chat text, which is then followed by a combination of features. Then, the features are sent to the LSTM to study the characteristics of the feature representation. Finally, the classification results are outputted. Experimental results show that the convolutional-LSTM model is better than the best baseline method by two percentage points on the F-measure.

Keywords travel consumption intention recognition, machine learning, deep learning, convolution neural network, long short-term memory neural network

Yue QIAN was born in 1992. He is a graduate student at the Harbin Institute of Technology, majoring in computer science and technology. His research interests include social computing and natural language processing.

Xiao DING was born in 1985. He received his Ph.D. from the Harbin Institute of Technology, Harbin, in 2016. He is currently an instructor at the Harbin Institute of Technology. His research interests include social computing and natural language processing.

Ting LIU was born in 1972. Currently, he is a professor and Ph.D. supervisor at the School of Computer Science and Technology, Harbin Institute of Technology. He is a senior member of CCF. His research interests include natural language processing, information retrieval, and social computing.

Yiheng CHEN was born in 1979. He is currently a lecturer at the Harbin Institute of Technology. His current research interests include information retrieval, social computing, and text clustering.