

基于篇章层次结构的商品评论摘要

张宜飞 王中卿 王红玲

苏州大学计算机科学与技术学院 江苏 苏州 215006

(20175227078@stu.suda.edu.cn)



摘要 商品评论摘要是从一个商品的所有评论中抽取出一系列有序的能够代表评论广泛意见的句子作为该商品的综合评论。篇章层次结构分析旨在对篇章内部各个语义单元之间的层次结构和语义关系进行分析。由此可见,分析篇章层次结构有利于更加准确地判断篇章内各个语义单元的语义信息和重要程度,这对于抽取篇章的重要内容有很大帮助。因此,文中提出了一种基于篇章层次结构的商品评论摘要方法。该方法基于 LSTM(Long Short Term Memory Network)神经网络构建抽取式商品评论摘要模型,并利用注意力机制将篇章层次结构信息作为判断篇章单元重要程度的参照加入该模型中,以便更加准确地抽取商品评论中的重要内容,从而提升整个任务的性能。将所提方法在 Yelp 2013 数据集上进行实验,并在 ROUGE 评价指标上进行评测。实验结果表明,加入篇章层次结构信息后,模型的 ROUGE-1 值达到了 0.3608,与仅考虑评论句子信息的标准 LSTM 方法相比提升了 1.57%,这说明在商品评论摘要任务中引入篇章层次结构信息能够有效地提升该任务的性能。

关键词: 商品评论摘要;篇章层次结构;神经网络;LSTM;注意力机制

中图法分类号 TP391

Product Review Summarization Using Discourse Hierarchical Structure

ZHANG Yi-fei, WANG Zhong-qing and WANG Hong-ling

School of Computer Science & Technology, Soochow University, Suzhou, Jiangsu 215006, China

Abstract Product review summarization aims to extract a series of relevant sentences that represent the overall opinions of the product. Analysis of discourse hierarchical structure aims to analyze the hierarchical structure and semantic relationship between the various semantic units in the discourse. Obviously, the analysis of discourse hierarchical structure is conducive to determine the semantic information and importance of each semantic unit in the discourse, which is very useful for extracting the important content of the discourse. Therefore, this paper proposed a product review summarization method based on discourse hierarchical structure. This method builds a product review summarization model based on LSTM and applies attention mechanism to extract the important content in the product review by integrating discourse hierarchical structure into the model. The experiments was conducted on the Yelp 2013 dataset and evaluated on the ROUGE evaluation index. The experimental results show that the ROUGE-1 value of the model after adding the discourse hierarchical structure is 0.3608, which is 1.57% higher than the standard LSTM method using only sentences information of the product review. This shows that the introduction of discourse hierarchical structure into the product review summarization task can effectively improve the performance of the task.

Keywords Product review summarization, Discourse hierarchical structure, Neural network, LSTM, Attention mechanism

1 引言

近年来,随着电子商务市场的发展,网上购物成为人们主要的购物方式。大量的电子商务网站,例如 Yelp, Amazon 等,通常不仅仅是一个产品展示和在线交易平台,还是一个允许用户自由发表评论的交流平台。研究发现,用户评论往往反映消费者对商品质量、功能和性能的主观体验,常常被潜在消费者参考^[1]。然而,随着评论数量的增加,一件热门商品常

常拥有成百上千条评论,这对于即将进行网上购物的用户来说,以人工分析的方式从大量评论中获取商品各个方面的综合评价变得十分困难。因此,如何自动地对商品的所有评论信息进行汇总,然后生成简洁且全面的摘要信息变得十分重要。为了解决这一问题,文中提出了商品评论摘要。

商品评论摘要的定义是从一个商品的所有评论中抽取一系列有序的能够代表评论的广泛意见的句子作为该商品的综合评论^[2-3]。也就是说,商品评论摘要可看作抽取式自动文

到稿日期:2018-12-15 返修日期:2019-04-13 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家自然科学基金青年科学基金项目(61806137,61702518);江苏省高等学校自然科学基金面上项目 48KJB520043)

This work was supported by the Young Scientists Fund of the National Natural Science Foundation of China (61806137,61702518) and Natural Science Foundation of the Jiangsu Higher Education Institutions of China (18KJB520043).

通信作者:王红玲(hlwang@suda.edu.cn)

摘^[4]的一种,其主要任务是根据一个商品评论中各个句子的重要性,挑选出最能代表评论核心内容的句子构成评论摘要。另外,一个商品的所有评论可以看作一个篇章,该商品的每一条评论可以看作篇章中的一个段落,评论中的每一个句子可以看作组成段落的句子。篇章(商品评论)内部存在着由单词构成句子、句子构成段落(评论)、段落(评论)构成篇章(商品评论)的层次结构。从语言学的角度来看,篇章内部各层语言单位之间的信息层层递进,相互影响,相同的句子在不同的段落或篇章中的语义和重要程度大不相同。因此,篇章的层次结构信息对抽取篇章摘要内容有很大帮助。

综上所述,本文提出了一个基于篇章层次结构的商品评论摘要方法。该方法基于标准 LSTM^[5] 构建抽取式摘要模型,在考虑句子信息的基础上加入篇章层次结构信息来抽取商品评论摘要。该方法主要包含以下 3 个步骤:首先,按照篇章层次结构将商品评论表示成“句子—评论—商品评论”的层次结构,并用 LSTM 神经网络来表示句子信息、评论信息和商品评论信息;然后,基于注意力机制(Attention)^[6] 对句子本身的内容进行强调,并形成新的句子信息、评论相关信息和商品评论相关信息;最后,基于篇章层次结构,将评论相关信息和商品评论相关信息作为预测句子重要程度的参照添加到句子信息中,即用向量[句子信息,评论相关信息,商品评论相关信息]作为预测句子重要度的输入。该输入经过一个全连接层之后得到句子是否为摘要的概率分布,本文选择概率最大的句子组合作为商品评论摘要。

本文第 2 节主要介绍商品评论摘要和篇章结构分析的相关工作;第 3 节阐述提出的方法;第 4 节介绍实验设置并分析实验结果;最后对本文研究工作进行总结,同时提出了下一步的研究方向。

2 相关工作

2.1 商品评论摘要

近年来,随着电子商务的发展,商品评论摘要任务已经成为自然语言处理领域的一个重要的研究方向。商品评论摘要不仅可以帮助用户更加快速、准确地了解商品的信息,从而做出正确的购买决定,而且可以帮助生产厂商了解商品的市场反响,从而做出合理的调整。常见的商品评论摘要方法通常包括以下 3 个步骤:商品特征抽取、情感分析/意见抽取和摘要生成^[7-8],即运用特征抽取和情感分析的结果指导生成商品评论摘要。2004 年,Hu 等^[7]提出了许多基于数据挖掘和自然语言处理的产品评论挖掘意见特征技术,然后使用特征抽取的结果选择句子以生成产品评论摘要。2010 年,Nishikawa 等^[9]为每个句子分配可读性和信息性两种得分,然后根据得分大小运用动态规划算法选择得分最高的句子作为商品评论摘要。同年,Ganesan 等^[10]提出了一种基于图模型的商品评论摘要生成方法,该方法将评论信息表示成图结构,通过规则方法反复探索图中各个子路径,从而选择最能表示图中信息的句子作为摘要。2012 年,Ganesan 等^[11]进一步采用一些启发式规则抽取商品评论中的重要短语,然后组合短语生成商品评论摘要,在此,他们使用修改的互信息函数和 n-gram 语言模型来确保抽取的短语的代表性和可读性。2014 年,

Gerani 等^[12]提出了基于篇章结构信息的评论摘要方法,该方法将评论信息按照其篇章结构表示成篇章结构图,然后采用基于模板的自然语言生成框架生成评论摘要。

2.2 篇章结构分析

篇章结构分析旨在分析篇章中各个语义单元之间的层次结构和语义关系。分析篇章结构有助于更加准确地判断篇章内各个语义单元的语义信息和重要程度,其对于提高自动文摘、知识抽取和自动问答等相关系统的性能均有重要作用^[13]。

近年来,随着篇章结构及其标引算法^[14]的研究和发展,很多研究者已经意识到篇章结构对自然语言处理任务的影响。2011 年,杨江等^[15]利用浅层篇章结构进行评论文倾向性分析,该方法利用情感主题句表示评论文的浅层篇章结构,避免了对全部篇章结构进行分析,进而排除了与主题无关的主观性信息。2015 年,Tang 等^[16]将篇章层次结构应用到情感分类任务中,以提高分类的准确度,并在实验数据集上取得了最优性能。2016 年,Yang 等^[17]将篇章层次结构应用到文本分类任务中,并在 6 个数据集上进行实验,实验结果表明,篇章层次结构信息能够极大地提高文本分类任务的性能。2017 年,Ren 等^[18]将篇章结构信息应用到摘要抽取任务中。他们对句子包含的词向量加权求和生成句子向量;然后,利用句子的语境信息,即上下文信息,对句子进行打分,作为该句重要性的依据。

受到上述工作的启发,本文将篇章层次结构信息作为区分篇章单元重要程度的参照加入商品评论摘要任务中,以便更加准确地评估商品评论中各个句子的重要性,从而提高该任务的性能。

3 问题分析和方法描述

摘要抽取的一个关键步骤是判断篇章内哪些词语或者句子更具有代表性,更能代表篇章内容。传统的神经网络模型是将文档进行分句,用句子内所有的单词信息构建句子的向量表示;然后将句子向量输入神经网络模型,以预测句子是否为摘要句的概率得分;最后系统按照句子得分从高到低地选择句子作为摘要^[19]。这种方法只考虑了句子本身的信息对句子重要程度的影响,而忽略了篇章内部段落信息和篇章信息等篇章层次结构信息对句子重要程度的影响。因此,本文提出将篇章层次结构信息加入商品评论摘要任务中,从而提高摘要抽取的效果。本文模型主要包含以下 4 个部分:篇章层次结构信息的选择、商品评论信息的层次化表示、句子打分和摘要生成。

3.1 篇章层次结构信息的选择

篇章是指由一系列连续的子句、句子或语段构成的语言单位,其内部存在着由单词构成句子、句子构成段落、段落构成篇章的层次结构。篇章中的各层语言单位之间的信息层层递进,相互影响,共同构成篇章整体。因此,篇章内句子的重要程度与其所在的段落内容和篇章内容密切相关。下面举例介绍不同篇章层次结构信息对句子重要性的影响。

3.1.1 段落信息

如表 1 所列,句子 1“环境实在太一般”同时出现在两个

不同的段落(评论)中。段落一中内容存在转折关系,段落重点强调“服务很好”,而段落二主要围绕句子1进行强调说明。通过分析可以看出:句子1在段落二中的重要程度明显高于在段落一中的重要程度。由此可见,句子的重要程度高度依赖于其所属段落信息的内容,仅在单个句子层面上很难准确地判断篇章内句子的重要程度。因此,在判断句子重要程度时加入段落信息应该是有效的。

表1 段落信息对商品评论摘要的影响示例

Table 1 Example of impact of paragraph information on product review summarization

句子1:“环境实在太一般”	
段落	段落一:昨天入住这家酒店,环境实在太一般。但是服务很好,有东西落在前台,给保管好并通知我们,很贴心,很贴心!!!
段落	段落二:环境实在太一般,进屋一股潮气扑来,整个屋子里都是湿湿的,设备也老旧,我对住宿要求很低了,但是对这酒店也是难以忍受。

3.1.2 篇章信息

表2是一个酒店的评价示例,该酒店评论可以作一个篇章,其包含的每一条评论可以看作篇章内的每一个段落。句子2“房间稍微有点小”是段落一中的一个句子,在这一段落中“服务不错”“环境一般”“房间稍微有点小”几部分内容存在并列关系,仅仅在段落层面上很难对这几个句子的重要程度进行判断。然而,在篇章层面上,由于篇章内各个段落都强调了“房间小”这一内容,句子2与篇章的主题思想相符,因此在这一篇章内,句子2的重要程度要高于段落一内其他句子的重要程度。由此可见,篇章信息对篇章内句子的重要程度有很大影响。

表2 篇章信息对商品评论摘要的影响示例

Table 2 Example of impact of document information on product review summarization

句子2:“房间稍微有点小”	
篇章	段落一:服务不错,有东西落在前台,给保管好并通知我们,很贴心。环境一般,靠路边吵,房间稍微有点小,其他还行吧。
	段落二:房间特小,卫生也不好,服务一般,不推荐!
	段落三:昨晚入住,房间很小,装修很陈旧,热水很小,空调制冷效果差,整晚冷死了。
	段落四:地方很好,就在景点附近。就是房间大小,电视机是坏的,前台服务态度也不怎么好。
	段落五:环境不好,房间很小,在一楼,有霉味,不好!!!

综上,本文选取了篇章层次结构信息中的段落信息和篇章信息进行实验,并评估了不同信息对句子重要程度的影响,从而找出最优的商品评论摘要模型。

3.2 商品评论信息的层次化表示

本文将一个商品的全部评论信息看成一个篇章,商品评论中的每条评论看成该篇章的段落,评论中的每个句子看成组成段落的句子。本文将商品评论信息按照其篇章层次结构表示成图1所示的层次结构。

对于商品评论信息 p_k (篇章),按照其包含的评论信息可分为评论集 $\{r_1, \dots, r_i, \dots, r_m\}$;对于任意一条评论 r_i (段落),按照其包含的句子信息可分为句子集 $\{s_{i1}, \dots, s_{ij}, \dots, s_{in}\}$ (句子)。由此,把商品评论信息表示成“商品评论-评论-句子”的层次结构,对应于篇章中“篇章-段落-句子”的层次结构。

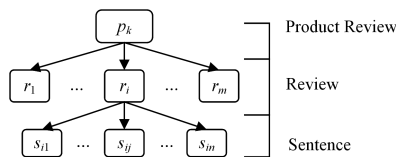


图1 商品评论信息的篇章层次结构

Fig. 1 Hierarchical structure of product review information

3.3 句子打分

本文使用基于篇章层次结构的商品评论摘要模型来给商品评论中的句子进行打分,从而决定文摘句的选择。将句子的分值作为句子的权重,权重越大,该句就越有可能是文摘句。该模型如图2所示,主要由以下6个部分组成。

1) 句子表示。对于商品评论中每一条评论 r_i 中的句子 s_{ij} ,首先,由句子内所有单词信息构建句子向量 $s = \{\omega_1, \omega_2, \dots, \omega_n\}$;然后,将该句子向量 s 输入LSTM^[5]神经网络模型中,得到相应的隐藏神经元向量 H_s ,并将 H_s 作为句子 s_{ij} 的表示。其中,句子向量 s 中的每一个单词 $\omega_k \in V$,表示句子 s_{ij} 中第 k 个单词在全局词表 V 中的位置,全局词表 V 为训练语料中包含的所有单词集合。

2) 评论表示。对于商品评论中的每一条评论,首先,由该评论包含的所有单词信息构成评论向量 r ;然后,把这个评论向量 r 输入LSTM神经网络模型中,得到相应的隐藏神经元向量 H_r 。本文将 H_r 作为评论 r_i 的表示。

3) 商品评论表示。对于商品评论 p_k ,首先,由该商品评论包含的所有单词信息构成商品评论向量 p ;然后,把这个向量 p 输入LSTM神经网络模型中,得到相应的隐藏神经元向量 H_p 。本文将 H_p 作为商品评论 p_k 的表示。

4) 基于评论信息(段落)的注意力机制。用句子的向量表示 H_s 为该句所属的评论向量 H_r 添加注意力机制^[6],从而对评论信息 H_r 中句子 H_s 部分的内容给予更大的关注度,如图2中边①和边②的指向所示,最后,得到评论相关向量的抽象表示 V_r 。其中,Attention机制的计算公式如下:

$$V_r = \sum_{j=1}^n a_j h_j \quad (1)$$

$$a_j = \frac{\exp(\omega_j)}{\sum_{j=1}^n \exp(\omega_j)} \quad (2)$$

$$\omega_j = \tanh(W^T h_j + b) \quad (3)$$

其中, h_j 表示评论向量 H_r 中的第 j 个单词的向量表示, a_j 表示句子向量 H_s 到单词 h_j 的注意力权重, \tanh 是一种非线性激活函数。 W 是中间矩阵, b 是偏移量,它们都在训练期间被随机初始化和更新。

5) 基于商品评论信息(篇章)的注意力机制。用句子的向量表示 H_s 为该句所属的商品评论向量 H_p 添加注意力机制,从而对商品评论信息 H_p 中句子 H_s 部分的内容给予更大的关注,如图2中边③和边④的指向所示,最后得到商品评论相关向量的抽象表示 V_p 。

6) 训练过程。首先,用句子的向量表示 H_s 为句子本身内容 H_s 添加注意力机制得到新的句子向量 V_s 。然后,将新的句子向量 V_s 、评论相关向量 V_r 和商品评论相关向量 V_p 进行拼接,如图2中边⑤、边⑥和边⑦的指向所示,得到最终预

测句子重要程度的输入。该输入经过一个全连接层之后得到句子 s_{ij} 是否为摘要的概率分布 $[P_0, P_1]$, P_0 表示该句子不是摘要句的概率, P_1 表示该句子是摘要句的概率。本文将 P_1 作为句子 s_{ij} 的得分, 并将其作为该句是否为摘要句的依据。

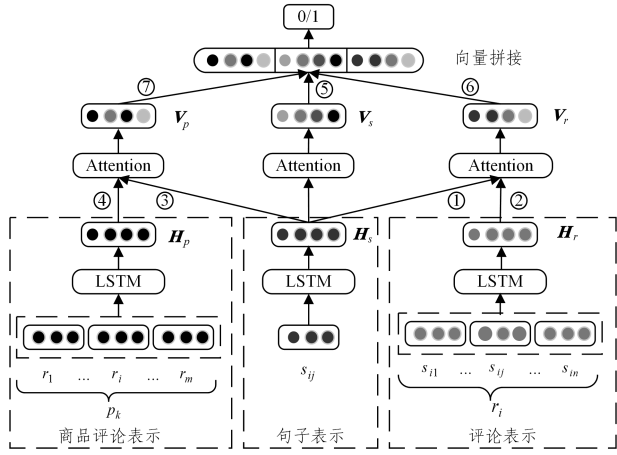


图2 基于篇章层次结构的商品评论摘要模型

Fig. 2 Overall architecture of product review summarization model using discourse hierarchical structure

3.4 摘要生成

本文使用句子抽取的方法生成摘要, 即先根据本文提出的基于篇章层次结构的商品评论摘要模型得到商品评论中各句子的得分, 然后系统按照句子得分从高到低地抽取句子组合作为摘要, 直到达到文摘指定的长度。

4 实验与评价

4.1 实验数据

本文使用美国著名的点评网站 Yelp2013¹⁾ 数据集对本文提出的模型的有效性进行验证。该数据集共包含 15 584 个 Business(商户)对象和 334 997 个 Review(评论)对象。其中, 商户对象与评论对象之间存在着一对多关系, 即一个商户对象包含许多评论对象。每种对象各自包含很多属性, 具体介绍如下。

Business: 一些餐馆或商店对象, 包含多条用户评论、该商户的评分和商户地址等相关信息。

Review: 评论对象, 包含该评论的文本信息、其从属的商户 ID 信息、该评论对应的评分信息以及该评论收到的投票数。其中, 投票数包含 funny, useful 和 cool 3 种。

本文以商户对象为单位, 选择评论数大于 100 的商户进行实验。最终, 我们选择了 715 家商户对象进行实验, 对每个商户只保留 10 条评论以及额外的 funny, useful 和 cool 投票总数最多的那条评论。本文随机抽取了 615 家商户, 共 6 150 条评论, 53 902 条句子作为训练集; 抽取了 100 家商户, 共 1 000 条评论, 8 575 条句子作为测试集。

4.2 评价方法

传统的自动摘要评价方法主要由人根据概述性、客观性、文法合理性、可读性以及内容含量等指标来判断文摘的质量。但是, 在对大规模文本进行人工评测时, 需要消耗大量的人

力, 实现起来比较困难。近年来, 如何进行文摘的自动评测引起了人们的极大重视。Lin 和 Hovy 提出的 ROUGE(Recall-Oriented Understudy for Gisting Evaluation)^[20] 方法是近几年摘要评价技术的通用标注之一。本文采用 ROUGE 中的 ROUGE-1, ROUGE-2, ROUGE-L 和 ROUGE-SU4 的 F 值来对实验抽取出的摘要进行评价。

4.3 对比实验

实验过程中, 为了分析不同篇章层次结构信息对商品评论摘要任务的影响, 文中分别构建了以下 6 个模型。

Random: 从商户的所有评论句子中随机抽取句子组合作为摘要, 直到达到文摘指定长度。

PageRank: 基于图排序的自动摘要算法。该方法把一个商户评论中的所有句子视作一个图结构, 将评论中的单个句子视为图中的顶点, 任意两句子之间的余弦相似度表示这两个顶点之间边的权重^[21]; 然后, 根据图中顶点之间的关系来确定图中各个句子的重要度; 最后, 按照句子重要度从高到低地抽取句子组合作为摘要, 直到达到文摘指定的长度。

标准 LSTM: 使用 LSTM 神经网络作为商户评论摘要的抽取方法。对于一个商户评论中的每一条句子 s_{ij} , 首先, 用该句子包含的单词信息构成该句子的向量表示 s ; 然后, 把句子向量 s 输入 LSTM 层进行特征提取, 得到句子隐藏神经元向量 H ; 最后, 将向量 H 输入全连接层, 预测出该句是否为摘要的概率分布。本文按照句子得分从高到低的顺序选择句子组合作为摘要, 直到达到文摘指定的长度。

LSTM+Review: 在上述标准 LSTM 方法的基础上添加了句子所在评论的信息, 即段落信息。首先, 用 LSTM 神经网络对句子和句子所在评论的信息进行表示; 然后, 用向量 [句子信息, 评论信息] 作为预测句子重要度的输入。

LSTM+Product: 在上述标准 LSTM 方法的基础上添加了句子所在商户评论的信息, 即篇章信息。首先, 用 LSTM 神经网络对句子和句子所在商户评论的信息进行表示; 然后, 用向量 [句子信息, 商户评论信息] 作为预测句子重要度的输入。

LSTM+Review+Product: 在上述标准 LSTM 方法的基础上添加了句子所在评论信息(段落信息)和商户评论的信息(篇章信息)。首先, 用 LSTM 神经网络对句子、句子所在评论信息和句子所在商户评论信息进行表示; 然后, 用向量 [句子信息, 评论信息, 商户评论信息] 作为预测句子重要度的输入。

4.4 实验设置

4.4.1 数据预处理

在实验中, 将每个商户评论中 funny, useful 和 cool 投票总数最多的那条评论作为该商户的理想评论, 并对每个商户做如下处理: 对该商户评论中所有句子按照句号、问号和感叹号等分隔符进行分句, 并以商户为单位构建“商户评论-评论-句子”的层次结构。在此, 本文限定每条句子包含的单词数在 4~50 之间, 从而过滤掉过长或过短的句子。对于训练集中的每一个商户对象, 计算该商户评论中每一个句子与该商户理想评论之间的相似度, 并将相似度最高的 10% 的句子标记

¹⁾ <https://www.yelp.com/dataset/challenge>

为摘要句,将剩余 90% 的句子标记为非摘要句。在此,本文选择句子向量之间的余弦值作为句子之间的相似度。

4.4.2 实验参数设置

本文采用监督学习^[22]方法进行商户评论摘要的抽取,首先,将训练集中商户评论中的句子结合其篇章层次结构信息输入本文提出的模型中,来训练句子分类模型;然后,将测试集中商户评论中的句子结合其篇章层次结构信息输入分类模型中,来预测该句子被分为摘要句的概率得分;最后,系统按照得分从高到低的顺序选择句子组合作为摘要,直到达到摘要指定长度。

根据语料统计得到训练集中所有商户的理想评论的平均长度为 283 个英文单词,因此本文将最终抽取的评论摘要的最大长度设置为 283。实验的详细参数设置如表 3 所列。

表 3 实验参数设置

Table 3 Setting of experimental parameters

参数	取值
EMBED_SIZE	64
HIDDEN_SIZE	128
MAX_LEN	50
BATCH_SIZE	300
EPOCHS	35

4.5 实验结果与分析

本文主要进行了 3 组实验来进行商品评论摘要的抽取。实验 1 主要用于分析不同篇章层次结构对实验结果的影响;实验 2 主要用于分析注意力机制对实验结果的影响;实验 3 主要是与一些基准模型进行比较,以衡量本文提出的基于篇章层次结构的摘要抽取模型的有效性。本文将实验生成的商品评论摘要与该商户的理想评论进行对比,并采用 ROUGE 评价指标进行结果评测。

4.5.1 不同层次结构信息的影响

为了分析篇章中不同层次结构信息对商品评论摘要任务的影响,本文采用了只包含句子本身信息(标准 LSTM)、加入句子所在评论信息(LSTM+Review)、加入句子所在商品评论信息(LSTM+Product)和加入句子所在评论信息和句子所在商品评论信息(LSTM+Review+Product)4 个模型进行对比分析,实验结果如表 4 所列。

表 4 增加不同层次结构信息的结果

Table 4 Results for adding different hierarchical information

方法	ROUGE-1	ROUGE-2	ROUGE-L	ROUGE-SU4
标准 LSTM	0.3451	0.0467	0.2274	0.1085
LSTM+Review	0.3532	0.0485	0.2292	0.1118
LSTM+Product	0.3539	0.0491	0.0231	0.1117
LSTM+Review+Product	0.3565	0.0511	0.2339	0.1150

从表 4 可以看出:首先,LSTM+Review 与标准 LSTM 相比,在 ROUGE-1 指标上提升了 0.81%,这说明评论(段落)有助于抽取篇章的重要内容;其次,LSTM+Product 与标准 LSTM 相比,在 ROUGE-1 指标上提升了 0.88%,这说明商品评论(篇章)信息对于抽取篇章的重要内容也有很大帮助;最后,LSTM+Review+Product 的实验结果达到最优,与标准 LSTM 相比,在 ROUGE-1 指标上提升了 1.14%。由此可见,

篇章中“句子-段落-篇章”之间的层次结构信息对句子本身的重要性有很大的影响,合理地运用篇章中不同层次的信息能够极大地提高摘要抽取任务的效果。

4.5.2 注意力机制的影响

为了分析注意力机制对模型有效性的影响,本文为上述 4 种方法(标准 LSTM、LSTM+Review、LSTM+Product 和 LSTM+Review+Product)分别添加注意力机制进行对比分析。实验结果如表 5 所列。

表 5 增加注意力机制的结果

Table 5 Results for adding attention mechanism

方法	ROUGE-1	ROUGE-2	ROUGE-L	ROUGE-SU4
标准 LSTM	0.3451	0.0467	0.2274	0.1085
LSTM+Attention	0.3513	0.0475	0.2309	0.1104
LSTM+Review	0.3532	0.0485	0.2292	0.1118
LSTM+Review+Attention	0.3556	0.0504	0.2318	0.1130
LSTM+Product	0.3539	0.0491	0.0231	0.1117
LSTM+Product+Attention	0.3567	0.0512	0.2328	0.1144
LSTM+Review+Product	0.3565	0.0511	0.2339	0.1150
LSTM+Review+Product+Attention	0.3608	0.0504	0.2341	0.1149

从表 5 可以看出:与标准 LSTM 相比,LSTM+Attention 在 ROUGE-1 指标上提升了 0.62%,其余 3 个模型在加入了注意力机制后,ROUGE 评价指标也均有一定的提升。这说明注意力机制能够对评论信息中重要的内容给予更大的关注度,从而提升整个模型的性能。

4.5.3 与基准系统的比较

为了衡量本文提出的模型的有效性,选择 Random 方法、PageRank 方法、标准 LSTM 摘要抽取方法与本文提出的模型进行了对比分析,实验结果如表 6 所列。其中,LSTM+Review+Product+Attention 是本文提出的基于篇章层次结构的模型,即在预测句子重要度时加入该句所属评论信息和商户评论信息,并用句子信息为评论信息和商户评论信息添加注意力机制。

表 6 对比实验结果

Table 6 Experimental results of comparative experiment

方法	ROUGE-1	ROUGE-2	ROUGE-L	ROUGE-SU4
Random	0.3422	0.0446	0.2227	0.1054
PageRank	0.3452	0.0519	0.2385	0.1142
标准 LSTM	0.3451	0.0467	0.2274	0.1085
LSTM+Review+Product+Attention	0.3608	0.0504	0.2341	0.1149

从表 6 可以看出:首先,PageRank 方法、标准 LSTM 方法和本文提出的 LSTM+Review+Product+Attention 方法的实验结果均高于 Random 方法,这说明本文选择的对比实验均是有效的;其次,本文所提模型的实验结果均高于对比方法,其中,所提模型的 ROUGE-1 值达到了 0.3608,与 Random 方法、PageRank 方法和标准 LSTM 方法相比,其在 ROUGE-1 指标上分别提升了 1.86%、1.56% 和 1.57%。这不仅说明本文提出的模型是有效的,也证明了篇章层次结构信息能够提升商品评论摘要抽取任务的效果。

结束语 篇章的层次结构决定了各个篇章单元之间相互联系、互相影响,这使得常见的摘要抽取模型很难准确地预测篇章中各个篇章单元的重要程度。因此,本文将篇章层次结构信息作为判断篇章单元重要程度的参照加入商品评论摘要抽取任务中,以便更加准确地抽取商品评论中相对重要的部分,从而提升整个任务的性能。实验结果表明:该方法与传统方法相比,在 ROUGE 值上均有较大的提升。这也证明了篇章层次结构信息有利于商品评论摘要任务,甚至对传统的抽取式摘要任务也有参考价值。在接下来的工作中,我们将进一步合理地利用篇章的结构信息,除了物理结构以外,还将引入如篇章修辞结构等语义结构,来更加准确地确定各个篇章单元之间的联系,从而提取出更高质量的评论摘要。

参考文献

- [1] KANG M, AHN J, LEE K, et al. Opinion mining using ensemble text hidden Markov models for text classification[J]. Expert Systems With Applications, 2017, 94.
- [2] LY D K, SUGIYAMA K, LIN Z, et al. Product review summarization from a deeper perspective[C]//ACM IEEE Joint Conference On Digital Libraries. ACM, 2011: 311-314.
- [3] YATANI K, NOVATI M, TRUSTY A, et al. Analysis of adjective-noun word pair extraction methods for online review summarization[C]//International Joint Conference on Artificial Intelligence. AAAI Press, 2011: 2771-2776.
- [4] ALLAHYARI M, POURIYEH S A, ASSEFI M, et al. Text Summarization Techniques: A Brief Survey[J]. International Journal of Advanced Computer Science and Applications, 2017, 8(10): 397-405.
- [5] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [6] SUN Q Y, WANG Z Q, ZHU Q M, et al. Stance Detection with Hierarchical Attention Network[C]//Proceedings of the 27th International Conference on Computational Linguistics. Santa Fe, New Mexico, USA, 2018: 20-26.
- [7] HU M, LIU B. Mining opinion features in customer reviews[C]//National Conference on Artificial Intelligence, American Association for Artificial Intelligence, 2004: 755-760.
- [8] HU M, LIU B. Mining and summarizing customer reviews[C]//Knowledge Discovery and Data Mining. ACM, 2004: 168-177.
- [9] NISHIKAWA H, HASEGAWA T, MATSUO Y, et al. Optimizing Informativeness and Readability for Sentiment Summarization[C]//Meeting of the Association for Computational Linguistics, 2010: 325-330.
- [10] GANESAN K, ZHAI C, HAN J, et al. Opinions: A Graph Based Approach to Abstractive Summarization of Highly Redundant Opinions[C]//International Conference on Computational Linguistics, 2010: 340-348.
- [11] GANESAN K, ZHAI C, VIEGAS E, et al. Micropinion generation: an unsupervised approach to generating ultra-concise summaries of opinions[C]//International World Wide Web Conferences. ACM, 2012: 869-878.
- [12] GERANI S, MEHDAD Y, CARENINI G, et al. Abstractive Summarization of Product Reviews Using Discourse Structure [C]//Empirical Methods in Natural Language Processing, 2014: 1602-1613.
- [13] LI Y C. Research of Chinese Discourse Structure Representation And Resource Construction[D]. Soochow: Soochow University, 2015.
- [14] SHAN Y M. Formal Analyses of Chinese Text Structure and Its Indexing Algorithm[C]//China National Conference on Computational Linguistics. Journal of Chinese information processing, 2001.
- [15] YANG J, HOU M, WANG N, et al. Sentiment Polarity Analysis of Reviews Based on Shallow Text Structure[J]. Journal of Chinese Information Processing, 2011, 25(2): 83-89.
- [16] TANG D, QIN B, LIU T, et al. Document Modeling with Gated Recurrent Neural Network for Sentiment Classification[C]//Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 2015: 1422-1432.
- [17] YANG Z, YANG D, DYER C, et al. Hierarchical Attention Networks for Document Classification[C]//North American Chapter of the Association for Computational Linguistics. Association for Computational Linguistics, 2016: 1480-1489.
- [18] REN P, CHEN Z, REN Z, et al. Leveraging Contextual Sentence Relations for Extractive Summarization Using a Neural Attention Model[C]//International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2017: 95-104.
- [19] SINHA A, YADAV A, GAHLOT, et al. Extractive Text Summarization using Neural Networks[J]. arXiv: Computation and Language, 2018.
- [20] LIN C Y, HOVAY E. Automatic evaluation of summaries using N-gram co-occurrence statistics[C]//Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology. 2003: 71-78.
- [21] WAN X, YANG J. Multi-document summarization using cluster-based link analysis[C]//Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2008: 299-306.
- [22] PEI Y, YIN W, FAN Q, et al. A Supervised Aggregation Framework for Multi-Document Summarization [C]//International Conference on Computational Linguistics, 2012: 2225-2242.



ZHANG Yi-fei, born in 1995, postgraduate, is member of China Computer Federation (CCF). Her main research interests include natural language processing and product review summarization.



WANG Hong-ling, born in 1975, assistant professor, is member of China Computer Federation (CCF). Her main research interests include natural language processing and text summarization.