

文章编号: 1003-0077(2019)09-0107-08

基于文本和用户信息的在线评论质量检测

吴 璠,王中卿,周夏冰,李寿山,周国栋

(苏州大学 计算机科学与技术学院,江苏 苏州 215006)

摘 要: 随着互联网的迅速发展,越来越多的用户评论出现在社交网站上。面对迅速增长的评论数据,如何为阅读评论的消费者提供准确、真实的高质量评论就显得尤为重要。评论质量检测旨在判断在线评论的质量,在传统的研究中,文本信息通常独立地被用于预测评论质量。但是在社交媒体上,每个文本之间不是独立的,而是可以通过发表文本的作者与其他文本相关联,即同一个用户或相近的用户发表的评论质量具有一定的相似性。因此,为了更好的构建文本的表示和研究文本之间基于用户的关联,该文基于神经网络模型分别构建用户和文本的表示,同时,为了放大用户信息的作用,我们进一步将基于注意力机制的用户信息融合到文本中,从而提高文本评论质量检测的效果。在 Yelp 2013 数据集上进行实验的结果表明,该模型能有效地提高在线评论质量检测的性能。

关键词: 评论质量;用户表示;神经网络模型;注意力机制

中图分类号: TP391 **文献标识码:** A

Neural Review Quality Prediction via Textual and User Information

WU Fan, WANG Zhongqing, ZHOU Xiabing, LI Shoushan, ZHOU Guodong

(School of Computer Science and Technology, Soochow University, Suzhou, Jiangsu 215006, China)

Abstract: With the rapid development of the Internet, more and more user comments appear on social networking sites. Review quality prediction aims to judge the quality of online reviews. To better build the representation of the text and study the user-based association between the text, this paper proposes a review quality prediction method of constructing the user and text representations based on neural network model. To properly emphasize the role of user information, we further integrate user information based on attention mechanism into the text to improve the effect of review quality prediction. Experiments on Yelp 2013 dataset show that our model can effectively improve the performance of online review quality detection.

Keywords: review quality; user representation; neural network model; attention mechanism

0 引言

随着互联网技术的普及,网络购物市场也迅速发展起来,网络用户规模不断扩大,在线评论网站如 Yelp、Amazon 等也迅速增多。对于网络购物的潜在消费者而言,其他用户的在线评论可以帮助他们做出有效的购买决策,他们通过阅读其他已购买消费的用户在线评论信息或与其他用户交流购物感受可以对商品有更深层次的了解,进而做出有效的选择,买到心仪的产品。而对于商家而言,通过挖掘大

量的评论数据来了解该项产品的用户体验,发掘用户最希望提供和改善的功能,以对产品做出针对性的改进。商家合理运用这些在线评论,将会为企业带来一波生产率的增长和利润的提高。对于产生在线评论的第三方网站来说,也可以通过提供高质量的评论信息来帮助他们建立良好的口碑,获取更高的关注度。

但是随着在线评论数量的不断增长,评论质量良莠不齐,故并非所有的评论都具备参考价值。同时由于网络的匿名性,用户非面对面接触,沟通成本低,内部操作空间大,以及各大在线评论网站对评论

收稿日期: 2018-10-15 定稿日期: 2018-12-26

基金项目: 国家自然科学基金(61836007, 61806137, 61702518, 61702149); 江苏省高等学校自然科学研究面上项目(18KJB520043)

内容的低约束性,导致一些用户随意发表评论,或提交一些与商品无关的评论,极端情况下,有的不良商家会找人撰写虚假、刻意夸大商品功能的评论。甚至一些同行竞争对手,会故意撰写诋毁对方商品或服务的恶意评论,这些评论对消费者的参考价值极低,甚至会引导消费者做出不正确的决定。在这种情况下,我们迫切地需要一个自动化的方法来辅助识别在线评论的质量。因此,科学、高效地从海量在线评论中帮助消费者抽取对决策参考价值高的评论信息就是本研究的价值所在。

目前,国内外的相关学者已经对评论质量检测做了相关研究,并取得了显著的成果。但是,以往的研究都认为每条评论是独立的,对于每条评论的质量进行单独的检测。本文与以往的研究不同,在社交媒体上,每个文本之间不是独立的,而是可以通过发表文本的用户与其他文本相关联的。我们在 Yelp 2013 实验数据集中随机抽取了两个用户发表在社交媒体上的两条评论,表 1 给出了评论示例及该评论在社交媒体上的得分。

表 1 用户评论示例

评论	打分
用户 A	
<i>Best meatball sandwich ! Rivals my mommas. Worth it !</i> (最好的肉丸三明治! 和我的妈妈竞争。值得!)	0
<i>The servers and hosts always wonderful.</i> (服务器和主机总是很棒。)	2
用户 B	
<i>Great place to chill and have some beer/wine,the selection of beers is extensive. They have live music usually on the weekends with some awesome local talent.</i> (放松的好地方,喝点啤酒/葡萄酒,啤酒的种类繁多。他们通常在周末有一些很棒的本地人才的现场音乐表演。)	8
<i>Such as rolled tacos being ice cold in the middle,too much cheeses put on all the dishes ,and bad service. You made me feel so uncomfortable when I ordered my food.</i> (比如,卷玉米饼中间还是冰冷的,所有的菜上都放了太多的奶酪,我点食物的时候你让我感到很不舒服。)	6

从表 1 的例子可以看出,用户 A 在社交媒体上发表的评论文本信息较为单薄,只有简单的类似“最好”“很棒”等形容词,具有强烈的个人倾向性,对大多数阅读评论的消费者参考价值不大,评论的质量普遍偏低。而用户 B 发表的评论描述相对仔细,评论有针对性,例如,会具体到“周末有本地人才的音乐表演”“点单服务态度”等相关的描述,这种表述对阅读评论的消费者选择的可参考性较大,他们可以根据自己的爱好和需求进行选择。因此,在社交媒体上,阅读评论的消费者对用户 A 的评论打分普遍偏低,对用户 B 的评论打分普遍偏高。这说明同一个用户或相近用户,在社交媒体上发表的评论质量具有一定相关性。

在本文的研究中,我们将同时考虑文本信息以及用户信息对于评论质量的影响。本文首先基于神经网络模型构建用户和文本的表示,同时为了研究用户信息对评论质量检测的影响,考虑到在线评论网站上的评论与传统评论文本的不同,社交媒体上,每个文本之间不是相互独立的,通过用户之间关联

可以找到这些评论质量之间的相关性。考虑到注意力机制在自然语言处理的其他任务上表现优异,因此,为了放大用户信息的作用,我们将基于注意力机制的用户信息融合到文本信息中。通过设计相关实验,验证在文本信息中加入针对用户信息的注意力机制,对评论的质量检测性能具有明显的提升作用。

本文组织结构安排如下:第 1 节介绍了情感分类和在线评论质量检测的相关工作;第 2 节介绍了数据收集的过程,且重点描述了本文构建的模型;第 3 节是实验设置的介绍和实验结果分析;最后,是对研究工作进总结,同时提出下一步的研究方向。

1 相关工作

评论的质量应该是以该条评论给消费者购物带来的参考价值来衡量的,评论给消费者带来的参考价值越高,评论质量就越高。所以,有很多在线评论网站会进行“有用性投票”,即通过一条评论所获得

的有用投票数占总投票数的比例来定义它的质量。实际上这种方法存在一定的弊端,可能会淹没一些高质量的评论。目前,对在线评论质量检测方面的相关研究主要从以下方面展开。

1.1 情感分类

情感分类一直是自然语言处理领域的研究热点,该任务的目标是对于给定的文本,推测其对应的情感极性,如 Positive、Negative 和 Neutral。情感分类的方法也有很多,有传统的基于支持向量机、朴素贝叶斯的分类方法,也有基于热门的深度学习分类方法。Wang 等^[1]提出多项朴素贝叶斯(MNB)模型和使用朴素贝叶斯特征的支持向量机模型(NBSVM)来进行情感分类。Pang 等^[2]以电影评论作为数据集,采用了三种机器学习方法:朴素贝叶斯、最大熵分类和支持向量机,实验表明这三种机器学习方法在情感分类中的表现不如在主题分类中。Turney^[3]提出了一种基于点互信息值来分析特定短语的情感极性,进而判断整篇文档情感倾向性的方法,该方法首先将文本进行分词和词性标注并提取出形容词或副词短语,然后使用“excellent”和“poor”两个种子词与未知词在搜索网页中的互信息来计算未知词的情感极性,并用以计算整个文本的情感极性。近年来,随着深度学习、神经网络的快速发展,很多神经网络技术也被应用到情感分类中,Cao 等^[4]发现卷积神经网络(CNN)与支持向量机(SVM)分类器相比,虽然情感分类的准确度得到了提升,但是不能有效地执行非线性分类,因此,他们提出了一个将 CNN 与 SVM 结合起来的模型,实验表明,该模型能达到较高的情感分类准确率。长短时记忆网络(LSTM)作为一种具有更复杂计算单元的递归神经网络,因其优良的序列信息保存能力,在各种序列建模任务中取得了很好的效果,Socher 等^[5]将 LSTM 模型应用到情感分类任务中,也取得了不错的效果。Tai 等^[6]将传统的 LSTM 结构推广到树状网络拓扑结构来进行情感分类。Qian 等^[7]通过损失函数将语言学规则引入到现有的句子级别情感分析的 LSTM 模型中,在没有增加模型复杂度的情况下,有效地利用情感词典、否定词和程度副词的信息,在实验数据集上取得了较好的效果。近年来,对情感分类的研究热度也一直不减,Songpan 等^[8]提出了一种新的情感分析方法—情感短语模式匹配(sentiment phrase pattern matching, SPPM),SPPM 与其他算法相比较,准确率、召回率和 F1 值

都有较大幅度的提高,该方法可以根据学生的意见来改善教学策略。Shen 等^[9]提出了一种基于扩展特征和动态合并的双通道卷积神经网络的文本情感分类算法,该算法比传统的单通道卷积神经网络算法具有更好的分类效果。但是以上研究都没有考虑评论质量的影响,低质量的评论会影响情感分类的准确率,因此,评论质量检测就显得尤为重要。

1.2 评论质量检测

在线评论质量检测研究,既可以看作为传统的文本分类问题,也可以看作成一个回归问题。以往的评论质量研究都主要集中在利用评论的文本信息,研究影响评论质量的不同特征。例如,Kim 等^[10]研究表明,评论发表距今的时间是显著影响评论质量的元数据特征。如 Ghose 等^[11]、Li 等^[12]、Liu 等^[13]指出影响评论质量检测主要的语言特征应包括评论字数、句子数、不同词性(名词、动词、形容词等)的词语数等。胡学钢等^[14]综合评论文本属性、评论者属性和店铺属性这三个影响因素,提出了一种基于多元线性回归算法的在线评论质量预测模型。Archak 等^[15]指出评论的极端性、评论的深度、评论的产品类型都将会对评论的质量造成影响。Chen 等^[16]强调在评论所包含的名词中,产品属性名词的频次是重要的语言特征,高质量的评论中应包含一定数量的产品属性名词。

近年来,也有一些关于发表评论的用户信息对评论质量检测影响的研究。Ghose 等^[11]认为,评论者相关信息是有效的评论质量检测特征,例如,评论者以往发表的评论数及有用率、评论者身份等。Cheng 等^[17]的研究侧重于分析评论者追随者人数、评论图片质量、评论字数以及周边路线这些社会因素对评论质量的影响。但是这些研究都是利用评论者自身的属性,例如,评论者好友的数量、注册的时间、发表的有用评论数。与上述利用评论者属性的研究不同,本文直接参考用户以往发表的在社交网站上的评论文本信息,利用同一个用户在社交网站上发表的评论质量具有一定的相似性,我们利用用户发表在社交网站上的历史评论文本信息构建用户的表示,以进一步提高评论质量检测的性能。

近几年来,注意力机制^[18-20]在自然语言处理的其他任务中表现优异。Yang 等^[21]针对文本分类提出一个层次化的注意力机制,分别就词和句子构建两个层次的注意力机制,从而使得文本中重要性不同的句子和词被给予了不同的“注意力”能力,实验

结果相比以往的模型效果有显著提升。Zhou 等^[22]基于注意力的双向 LSTM 来进行关系分类也取得了不错的效果。Sun 等^[23]利用层次化注意力机制权衡各种语言信息,将其应用到立场检测上也效果显著。本文受到启发,在进行评论质量检测分析模型构建时引入用户信息,并且,为了进一步放大用户信息的影响,我们将基于注意力机制的用户信息融合到文本中,经过在 Yeap.com^①数据集上测试,模型的实验效果得到稳步提升。

2 基于注意力机制用户信息的评论质量检测

本文利用评论的文本信息构建 LSTM 模型对评论的质量进行检测,考虑到在线评论网站上的评论与以往独立的单文本不同,社交网站上,每个评论文本之间不是相互独立的,通过用户之间相关性可

以找到这些评论质量之间的相关性。一般来说,拥有更多粉丝和已发表的高质量评论数越多的用户,再次发表的评论质量会更高,因为同一个用户或者相近的用户发表的评论质量也是相似的。因此,我们在模型中加入了用户信息,从而对评论的质量进行全面的预测。最近出现的注意力机制在其他自然语言处理任务上都取得了不错的效果,注意力机制可以更好地表征文本,在训练时可以自动获取更为重要的特征。例如,在情感分类任务中,地名、人名等名词没有表达情感类的形容词重要,所以,在神经网络训练时,表达情感的形容词就应该占有更大的权重。因此,引入注意力机制可以将神经网络的注意力集中在那些对当前任务更重要的向量上,以提高模型的准确率。所以,为了进一步放大用户信息的作用,我们结合注意力机制将用户信息融合到文本信息中。实验结果表明,评论质量检测分析的性得到了进一步提高,本文的整个模型如图 1 所示。

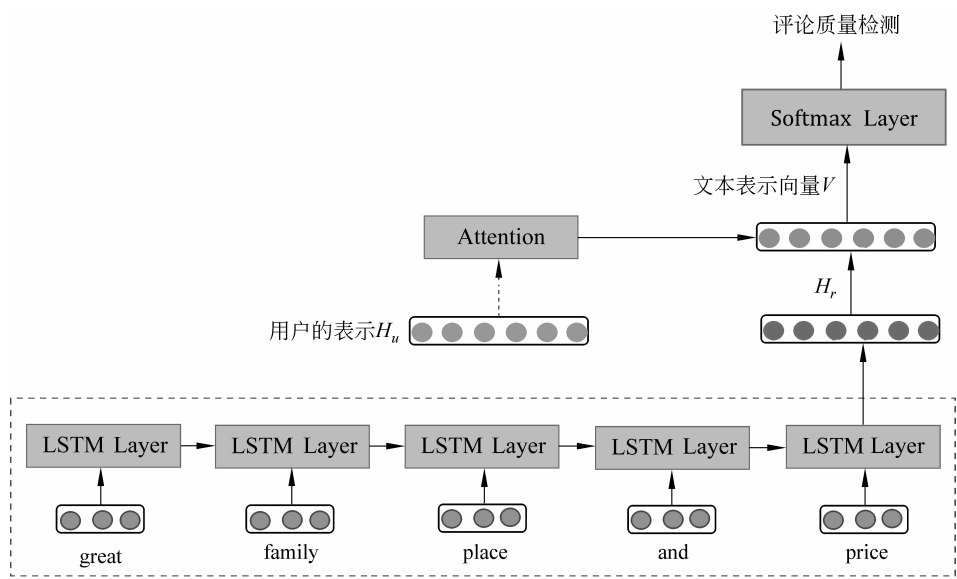


图 1 注意力机制模型图

2.1 文本信息的表示

在线评论是由用户针对特定产品做出的评价,我们将每条评论表示为由 $\{W_1, W_2, W_3, \dots, W_n\}$ 组成的 n 个单词。LSTM 长短时记忆网络是 RNN (Recurrent Neural Network) 中一种特殊的类型,所有的 RNN 模型都具有一种重复神经网络模块的链式形式。与传统的 RNN 模型不同,1997 年,Hochreiter 和 Schmidhuber 首先提出了 LSTM 神经单元。将传统 RNN 中的隐含层以 LSTM 单元替代,也就是现在所说的 LSTM 网络。对于每一条评论,我们

使用 LSTM 神经网络来学习评论的深层文本信息表示。首先,通过查找预训练好的词向量表将每个标记的 W_i 转换为实值向量 X_i 。然后,我们使用 LSTM 神经网络来生成隐藏的向量序列 $\{h_1, h_2, h_3, \dots, h_n\}$ 。在 LSTM 神经网络的步骤 t 中,LSTM 的隐藏向量 $h_t = \text{LSTM}(H_{t-1}, X_t)$ 均可由当前向量 X_t 和先前向量 h_{t-1} 来获得。LSTM 神经网络中的参数初始化都是随机初始化,在训练过程中调整的。

LSTM 单元利用记忆细胞 (Memory Cell) 来传

① www.yelp.com

递信息,因此细胞状态是整个 LSTM 的关键,它类似于信息传统带,直接在整个循环网络中运行,且由它来决定该保留或遗弃哪些信息。历史信息由 3 个门控制,即输入门 (Input Gate),遗忘门 (Forget Gate) 和输出门 (Output Gate)。LSTM 表示如式(1)~式(6)所示。

$$f_t = \sigma(W_f \cdot (H_{t-1}, X_t) + b_f) \quad (1)$$

$$i_t = \sigma(W_i \cdot (H_{t-1}, X_t) + b_i) \quad (2)$$

$$C_t = \tanh(W_c \cdot (H_{t-1}, X_t) + b_c) \quad (3)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (4)$$

$$O_t = (W_o \cdot (H_{t-1}, X_t) + b_o) \quad (5)$$

$$h_t = O_t \cdot \tanh(C_t) \quad (6)$$

其中, i_t, f_t, O_t , 和 C_t 分别表示 t 时刻对应的三种门结构和细胞状态。LSTM 是一种特殊的 RNN,是为了解决 RNN 中长期依赖问题而专门设计出来的,同时 LSTM 对于大多数自然语言处理任务也确实展示了更好的性能。采用 LSTM 模型避免了传统机器学习模型中的特征工程,而且能取得与当前基于传统机器学习相当甚至更好的性能。并且 LSTM 等深度学习模型还可以借助 GPU 来训练,大大缩短了训练时间。

2.2 用户信息的抽取

本文的研究目的是在线评论的评论质量。根据日常经验,很容易联想到发表评论用户的相关信息会对评论质量造成影响。拥有更多粉丝数,已发表的有用评论数量越多的用户,再次发表的评论质量可能会更高,而发表评论质量一直较低的用户,再次发表的评论质量较低的可能性也比较大。与以往直接利用评论者粉丝数,和评论者发表的评论总数等用户属性的研究不同,社交媒体网站上的评论之间并不是相互独立的,而是可以通过用户之间的联系将评论联系起来。因此,在实验中,我们考虑加入了用户在社交媒体上发表的历史评论信息,利用 LSTM 模型学习用户评论深层的信息作为用户的表示(图 2)。

在 LSTM 模型中,将历史评论信息 $\{e^d(r_{u_1}), e^d(r_{u_2}), e^d(r_{u_3}), \dots, e^d(r_{u_n})\}$ 作为输入,令 $h_{u_i} = \text{LSTM}(H_{u_{i-1}}, e^d(r_{u_i}))$,经过 LSTM 神经网络得到的隐藏状态向量 $\{h_{u_1}, h_{u_2}, h_{u_3}, \dots, h_{u_n}\}$,将其作为每个用户的表示。在具体实验中,我们对于每个用户,抽取其发表的 n 条评论,并利用 LSTM 模型基于这些评论构建每个用户的表示。

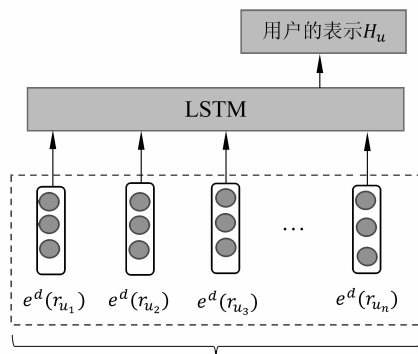


图 2 用户的表示示例图

2.3 基于注意力机制和用户信息的评论质量检测

注意力机制用于从众多信息中选择出对当前任务目标更关键的信息。而在该任务中并不是每一个评论都对用户的表示做出了同等贡献,因此,我们在获取了文本表示和用户表示之后,为了放大用户信息对评论质量的影响,针对用户信息引入了注意力机制,用于提取用户相对重要的评论,形成基于注意力机制新的用户表示。

注意力机制将每一个输入状态 $\{h_{u_1}, h_{u_2}, h_{u_3}, \dots, h_{u_n}\}$ 作为用户的表示输入,从而实现为输入文本构造一个隐藏状态 h_{u_i} 的加权和,输出引入用户注意力机制的文本表示向量 v ,如式(7)所示。

$$v = \sum_{i=1}^T \partial_i h_{u_i} \quad (7)$$

其中, T 为输入状态的数目, ∂_i 为输入状态 h_{u_i} 的权重, $\partial_i \in [0, 1]$, 并且 $\sum_i \partial_i = 1$, ∂_i 的计算如式(8)、式(9)所示。

$$\partial_i = \frac{\exp(u_i)}{\sum_{j=1}^T \exp(u_j)} \quad (8)$$

$$u_i = \tanh(W h_{u_i} + b) \quad (9)$$

其中, W 为模型中的权重, b 为模型中的偏置。

实验表明,通过针对用户信息引入注意力机制,使得实验结果与仅引入用户信息比较得到了进一步的提高。这说明加入基于用户信息的注意力机制确实能进一步提高评论质量检测的准确度。

3 实验

3.1 实验数据集

本文使用的是 Yelp Dataset Challenge 2013 提

供的数据。在实验中,选取随机生成的用户 ID 中前两千条用户 ID 的评论作为训练集,两千到四千条用户 ID 的评论作为测试集。表 2 反映了实验数据集的分布情况。

表 2 数据集分布情况

数据集	Yelp 2013
# class	5
# docs	78 966
# users	1 310
# products	1 633
# sens/doc	10. 89
# words/sen	17. 38

在 Yelp 数据集中: 每个评论至少包括以下信息。

(1) 每个评论会有一个 useful score,这是该条评论的质量指标,useful score 越高,表示该条评论的质量越高,参考价值越大。在实验中,为了防止极少数 useful score 极端值对实验结果的影响,在实验中将 useful score 的上限设置为 10,大于 10 的 useful score 看作 10 来处理,同时将 useful score 转化为 0 到 1 之间的数值,在评论质量检测时作为 useful label。

(2) 评论的文本信息。

(3) 评论的用户 ID 信息。

(4) 评论的星级。最高星级为五颗星,最低为一颗星。星级用来表示评论的情感倾向,星级越高,表示用户对该产品越满意,大于三颗星的评论即为积极情感倾向的评论。

(5) 评论的对象商家或产品的 ID 信息。

本文采用均方误差作为评论质量检测的评价指标,因为均方误差能很好地反映预测评论质量标签和真实评论质量标签的偏离程度,评价标准 MSE 的公式如式(10)所示。

$$MSE = \frac{\sum_i^N (gr_i - pr_i)^2}{N}$$

(10)

其中, N 为整个数据集评论数, gr_i 和 pr_i 分别为真实评论质量 (useful label) 和预测的评论质量 (useful label)。同时,为了加速训练,我们限定每条评论最多有 40 个句子,每个句子不超过 40 个单词。均方误差比较直观且计算简单,可以评价数据的变化程度,MSE 的值越小,说明预测模型的预测结果

具有更好的准确度。

3.2 实验参数设置

为了获取模型的最优参数集合,需要调节的参数如表 3 所示。

表 3 模型参数调节列表

参数	说明	取值
Optimizer	模型优化器	Adam
Learn_rate	学习率	0. 0006
Batch_size	梯度下降时,每个 batch 包含的样本数	20
dropout	每次更新参数时,随机断开输入神经元的比例,用于防止过拟合	在模型的输入层之后和最后的 softmax 层之前分别使用了 dropout,取值都为 0. 2

3.3 基准系统

在前人评论质量检测研究的基础上,我们选用以下方法进行实验的对比分析。

SVM 模型、Decision Tree 模型: Liu 等^[13]采用支持向量回归、决策树等机器学习方法进行评论质量检测比较,以得到性能最优的模型。本文以单个词 Unigram 作为特征,使用传统机器学习中 SVM^① (support vector machine)模型、决策树模型进行回归分析。

LSTM 模型: Bobby^[24]使用深度学习模型 LSTM 对亚马逊网站上的评论进行质量检测。该模型不引入用户信息和注意力机制,我们直接采用 LSTM 模型对评论文本进行评论质量检测,作为本文的基准模型。

GRU 模型: Cho 等^[25]提出了另一种循环神经网络的变体 GRU(Gate Recurrent Unit),它的实验效果与 LSTM 相似,但更易于计算。

Bidirectional-GRU 模型: 双向的 GRU 模型优于单向 GRU 是其可以同时利用过去时刻和未来时刻两个方向上的信息,从而进一步提高模型的准确度。

LSTM+User 模型: 该模型在 LSTM 模型的基础上加入了评论的用户信息,实现评论质量检测。

LSTM+SelfAttention 模型: 该模型在 LSTM 模型的基础上加入了自注意力机制,实现评论质量

① <http://svmlight.joachims.org/>

检测。

LSTM + User + Attention 模型：该模型在 LSTM 模型的基础上将基于注意力模型的用户信息融合文本中，相比以上的几种模型，实验性能得到进一步提升。

3.4 实验结果与分析

3.4.1 与传统方法对比

本文主要进行了两组实验。为了体现我们提出的结合用户信息注意力机制模型的有效性，本文的第一组实验选择与前人所使用的 SVM、决策树等传统机器学习进行对比，同时也与自然语言处理领域一直表现优异的长短时记忆网络 LSTM、GRU、双向 GRU 模型进行对比，实验结果如表 4 所示。

表 4 与 baseline 进行对比

对比实验	模型名称	MSE 值
Liu ^[13]	SVM	0.119
	Decision Tree	0.159
	LSTM	0.096
Cho ^[25]	GRU	0.098
	Bidirectional-GRU	0.090
	LSTM+User+Attention	0.082

从表 4 可以看出，本文提出的基于神经网络 LSTM、GRU 模型确实能学习到文本更深层次的信息，比传统的机器学习方法在评论质量检测中性能更好，在一定程度上体现了深度学习的优势。同时，对比前人使用的模型和我们提出的结合用户注意力机制的模型，发现加入用户信息注意力机制对评论质量检测模型性能确实有一定的提升。

3.4.2 网络结构的影响

为了验证模型的有效性，考虑不同网络结构的影响，研究用户信息注意力机制的作用，我们进行了第二组实验，实验结果如表 5 所示。

表 5 不同模型 MSE 值对比

模型名称	MSE 值
LSTM	0.096
LSTM+User	0.091
LSTM+SelfAttention	0.084
LSTM+User+Attention	0.082

从表 5 的实验结果来看，在 LSTM 模型中，直

接加入用户信息模型的性能得到了提高。同时，比较“LSTM + SelfAttention”和“LSTM + User + Attention”模型性能，我们发现加入针对用户信息的注意力机制，能进一步提高评论质量检测的性能。

通过对比实验可以看出，逐步加入用户信息和针对用户信息引入注意力机制可以降低 MSE 值，提高评论质量检测的性能。这表明，在评论质量检测中评论者以往的评论信息是影响评论质量的一个重要因素，通过将评论的用户信息和注意力机制相结合，有效增强了模型对于文本信息的捕获能力，进而提高了评论质量检测的性能。

4 结论

互联网上日益丰富的评论信息确实给人们的生活带来了便利，为人们在购物时的决策提供了更多的参考。但互联网上的信息良莠不齐，数据庞大，信息爆炸也为消费者带来了困惑，同时也影响了评论的参考价值，降低了消费者决策的效率和效果。因此快速地筛选出对消费者决策最有帮助的高质量评论就显得尤为重要。

考虑到社交媒体上同一个用户发表评论质量的相似性，本文提出一种基于 LSTM 神经网络，针对用户信息引入注意力机制来预测在线评论质量的模型。该模型能帮助消费者从在线评论网站海量的评论中快速识别出有用的评论，做出正确的购买决策。

类比在模型中加入评论的用户信息，我们很容易联想到，加入评论的对象商家信息是否也会提高评论质量分析的性能。我们在分析实验语料时发现，评论者的好友信息也具备一定的参考价值，有用评论的用户好友质量普遍比较高，发表的评论数比较多，同时有用的评论也比较多，因此，可以考虑在模型中加入用户的好友信息。同时，我们受到情感分类的启发，分析评论的情感信息或许也能更好地筛选质量高的评论。以上所述都有待未来更详尽的研究工作来完成，本文以后的工作将进一步考虑加入商家信息和情感信息，并探究其对评论质量的影响。

参考文献

[1] Wang S, Manning C D. Baselines and bigrams: simple, good sentiment and topic classification[C]//Proceedings of the 50th Annual Meeting of the Association for

- Computational Linguistics, 2012; 90-94.
- [2] Pang B, Lee L, Vaithyanathan S. Thumbs up? Sentiment classification using machine learning techniques [C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2002; 79-86.
 - [3] Turney P. Thumbs Up or Thumbs Down? Semantic orientation applied to unsupervised classification of reviews [C]//Proceedings of the Annual Meeting of the Association for Computational Linguistics, 2002; 79-86.
 - [4] Cao Y, Xu R, Chen T. Combining convolutional neural network and support vector machine for sentiment classification [C]//Proceedings of the Meeting of the Chinese National Conference on Social Media Processing, 2015; 144-155.
 - [5] Socher R, Perelygin A, Wu J Y, et al. Recursive deep models for semantic compositionality over a sentiment treebank [C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2013; 1631-1642.
 - [6] Tai K S, Socher R, Manning C D. Improved semantic representations from tree-structured long short-term memory networks [J]. Computer Science, 2015, 5(1): 36.
 - [7] Qian Q, Huang M, Lei J, et al. Linguistically regularized LSTM for sentiment classification [C]//Proceedings of the Annual Meeting of the Association for Computational Linguistics, 2017; 1679-1689.
 - [8] Pong-Inwong C, Songpan W. Sentiment analysis in teaching evaluations using sentiment phrase pattern matching (SPPM) based on association mining [J]. International Journal of Machine Learning & Cybernetics, 2018(6): 1-10.
 - [9] Shen C, Junzhong J I. Text sentiment classification algorithm based on double channel convolutional neural network [J]. Pattern Recognition & Artificial Intelligence, 2018, 31(2): 759-762.
 - [10] Kim S M, Pantel P, Chklovski T, et al. Automatically assessing review helpfulness [C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2006; 423-430.
 - [11] Ghose A, Ipeirotis P G. Estimating the helpfulness and economic impact of product reviews: Mining Text and Reviewer Characteristics [J]. IEEE Transactions on Knowledge and Data Engineering, 2011, 23(10): 1498-1512.
 - [12] Fan L, Zhang Y, Dang Y, et al. Analyzing sentiments in Web 2.0 social media data in Chinese: experiments on business and marketing related Chinese Web forums [J]. Information Technology and Management, 2013, 14(3): 231-242.
 - [13] Liu Y, Jin J, Ji P, et al. Identifying helpful online reviews: A product designer's perspective [J]. Computer-Aided Design, 2013, 45(2): 180-194.
 - [14] 胡学钢, 陈方鑫, 张玉红. 在线商品评论有用性影响因素研究 [J]. 计算机应用研究, 2016, 33(12): 3559-3561.
 - [15] Archak N, Ghose A, Ipeirotis P G. Deriving the pricing power of product features by mining consumer reviews [M]. Inform. 2011, 57(8): 56-65.
 - [16] Chen C C, Tseng Y D. Quality Evaluation of product reviews using an information quality framework [J]. Decision Support Systems, 2011, 50(4): 755-768.
 - [17] Cheng Y H, Ho H Y. Social influence's impact on reader perceptions of online reviews [J]. Journal of Business Research, 2015, 68(4): 883-887.
 - [18] Mnih V, Heess N, Graves A. Recurrent models of visual attention [C]//Proceedings of Advances in Neural Information Processing Systems, 2014(3): 2204-2212.
 - [19] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate [C]//Proceedings of the ICLR, 2015: 1-15.
 - [20] Luong M, Manning C D. Effective approaches to attention-based neural machine translation [C]//Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, 2015: 1412-1421.
 - [21] Yang Z, Yang D, Dyer C, et al. Hierarchical attention networks for document classification [C]//Proceedings of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2017: 1480-1489.
 - [22] Zhou P, Shi W, Tian J, et al. Attention-based bidirectional long short-term memory networks for relation classification [C]//Proceedings of Annual Meeting of the Association for Computational Linguistics, 2016: 207-212.
 - [23] Qingying S, Zhongqing Wang, Qiaoming Zhu, et al. Stance detection with hierarchical attention network [C]//Proceedings of the 27th International Conference on Computational Linguistics, 2018.
 - [24] Gulcehre C, Cho K, Pascanu R, et al. Learned-Norm pooling for deep feedforward and recurrent neural networks [C]//Proceedings of European Conference on Machine Learning and Knowledge Discovery in Databases, 2014: 530-546.
 - [25] Gulcehre C, Cho K, Pascanu R, et al. Learned-norm pooling for deep feedforward and recurrent neural networks [C]//Proceedings of the European Conference on Machine Learning & Knowledge Discovery in Databases, 2014: 530-546.

[24]

Koo T, Collins M. Efficient third-order dependency parsers[C]//Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics. 2010:1-11.

[25]

张庆庆, 刘西林. 基于依存句法关系的文本情感分类研究[J]. 计算机工程与应用, 2015, 51(22):28-32.

[26]

Golberg D E. Genetic algorithms in search, optimization, and machine learning[M]. Addison-Wesley, Boston, MA. 1989.

[27]

余有明, 刘玉树, 阎光伟. 遗传算法的编码理论与应用[J]. 计算机工程与应用, 2006, 42(3):86-89.

[28]

刘志明, 刘 鲁. 基于机器学习的中文微博情感分类实证研究[J]. 计算机工程与应用, 2012, 48(1):1-4.

[29]

杨立月, 王移芝. 微博情感分析的情感词典构造及分析方法研究[J/OL]. 计算机技术与发展, 2019(02): 1-6.



潘浩(1995—), 硕士研究生, 主要研究领域为情感分析与实体识别。
E-mail: panhao950810@sjtu.edu.cn



卫宇杰(1996—), 硕士研究生, 主要研究领域为质量舆情与质量标准化体系。
E-mail: jd-wyj@sjtu.edu.cn



潘尔顺(1972—), 通信作者, 博士, 教授, 主要研究领域为可靠性工程与宏观质量研究。
E-mail: pes@sjtu.edu.cn



(上接第 114 页)



吴瑶(1994—), 硕士研究生, 主要研究领域为自然语言处理。
E-mail: fwul23@stu.suda.edu.cn



王中卿(1987—), 通信作者, 博士, 讲师, 主要研究领域为自然语言处理, 社会计算。
E-mail: wangzq@suda.edu.cn



周夏冰(1988—), 博士, 讲师, 主要研究领域为自然语言处理。
E-mail: zhouxiabing@suda.edu.cn