

# 复合双通道酒店评论方面类别情感分析

李 薇, 王欣羽

燕山大学理学院, 河北 秦皇岛

收稿日期: 2023年9月30日; 录用日期: 2023年10月27日; 发布日期: 2023年11月3日

## 摘 要

针对中文酒店评论情感分析会忽视特定方面情感极性的问题, 提出一种复合双通道模型GCGAT, 进行方面类别情感分析。该模型基于门控机制和双向门控循环单元, 由四层网络结构组成; 利用门控机制控制通道输出; 引入方面类别影响因素词向量和注意力机制提高模型信息提取准确性; 数据集方面, 利用LDA-Jaccard模型进行聚类, 确定方面类别并进行标注, 建立了细粒度数据集。通过对比实验发现, GCGAT模型取得了很好的效果, 在准确率、F1值、ROC曲线和auc值上效果均优于其他模型。

## 关键词

卷积神经网络, 双向门控循环单元, 门控机制, 注意力机制, 方面类别情感分析

# Composite Dual-Channel Sentiment Analysis of Hotel Reviews in Terms of Categories

Wei Li, Xinyu Wang

School of Science, Yanshan University, Qinhuangdao Hebei

Received: Sep. 30<sup>th</sup>, 2023; accepted: Oct. 27<sup>th</sup>, 2023; published: Nov. 3<sup>rd</sup>, 2023

## Abstract

Aiming at the problem that the sentiment analysis of Chinese hotel reviews will ignore the polarity of specific aspectual sentiment, a composite two-channel model, GCGAT, is proposed for aspectual category sentiment analysis. The model is based on the gating mechanism and bidirectional gating loop unit, and consists of a four-layer network structure; the gating mechanism is utilized to control the channel output; the aspect category influencing factor word vector and the attention mechanism are introduced to improve the accuracy of the model's information extraction; in terms of the dataset, clustering is carried out by using the LDA-Jaccard model to determine the aspect categories and annotate them, and a fine-grained dataset is established. Through comparison expe-

riments, it was found that the GCGAT model achieved good results, and the effect was better than other models in terms of accuracy, F1 value, ROC curve and auc value.

## Keywords

CNN, biGRU, Gating Mechanism, Attention, Aspect Category Sentiment Analysis

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着互联网的发展,对在线评论进行情感分析逐渐显露巨大价值[1],研究方法从机器学习[2] [3]发展到深度学习[4] [5]和预训练模型[6] [7]。随着研究的深入,细粒度情感分析逐渐走入研究视野,方面类别情感分析作为细粒度情感分析中的一个重要子任务,到目前为止已取得广泛的研究关注[8] [9] [10]。旅游作为现如今人们放松身心的必要项目,近年来已取得飞速发展,根据国家统计局公布的数据显示,近十年我国住宿业营业额呈上涨趋势,在2021年我国住宿业营业额达到4071.59亿元[11]。对酒店在线评论情感分析,可进一步促进酒店行业发展,但目前对中文酒店评论进行情感分析的研究大多从粗粒度角度出发。邬小燕[12]通过改进Stacking中的SVM和KNN两个基分类器,以及改进TF-IDF加权方法,提高模型对酒店评论的情感极性分类效果。高华玲和张晶[13]将通用情感词典HowNet和酒店评论相关的评论领域专业词典进行结合,通过构建的领域情感词典对高端酒店评论进行情感分析并进行可视化展示,最终对高端酒店的经营策略给出可行建议。丁美荣等[14]提出两种方法对酒店评论进行情感分析,一种是基于扩展情感词典的方法,通过计算情感值进行分类,另一种是采用双向长短时记忆模型(Bi-directional Long Short-Term Memory, BiLSTM)对预训练数据集进行情感分析。

基于以上研究,针对酒店评论情感分析,本文提出进行方面类别情感分析,并利用所提出的GCGAT(GTU-CNN-GRU-Attention)模型进行分析,该模型由四层网络构成,分别为:特征提取层、语义信息提取层、特征信息权重分配层和最后的全连接层。其中,特征提取层利用卷积神经网络(Convolutional Neural Networks, CNN)和门控机制[15]对评论信息和嵌入方面类别的评论信息进行情感特征和方面类别特征提取;语义信息提取层利用双向门控循环单元(Bi-directional Gate Recurrent Unit, BiGRU)进行上下文信息的提取;特征信息权重分配层利用最大池化和注意力机制[16]引导权重分配;最终通过具有输出的全连接层进行情感极性判定。在数据集方面,本文使用数据采集软件获取多平台的酒店评论作为原始数据,通过LDA-Jaccard主题模型对数据进行主题分析并确定方面类别,根据确定的方面类别对数据进行人工标注。本文主要贡献如下:

- 1) 利用LDA-Jaccard主题模型对预处理后的数据进行分析确定方面类别,建立了细粒度酒店评论数据集;
- 2) 本文提出对酒店评论进行方面类别情感分析,并提出基于门控机制和BiGRU的深度学习模型,该模型通过四层网络结构对评论中的情感特征、方面类别特征和上下文信息进行充分提取,并合理分配权重,有效提高了情感极性判定准确率;
- 3) 细粒度酒店评论数据集上的实验结果表明,GCGAT模型性能要优于其他模型。

## 2. 相关工作

### 2.1. 情感分析

在早期的情感分类任务中, 研究学者主要运用基于统计的非深度学习方法进行研究, Kiritchenko [17] 通过使用基于 SVM 模型的词典、语法和 n-gram 特征方法进行情感分析, 并在 SemEval2014 数据集上取得了最佳性能。王新宇[18]对旅游平台的线上评论进行研究分析, 采用 TF-IDF 方法对各个特征分配权重, 最后使用 SVM 作为分类器, 对情感极性进行分类。由于 Mikolov 等[19]提出的 word2vec 模型提供了一种简单并有效的分布式表示方法, 使深度学习方法在情感分析中得到了广泛的应用。徐菲菲和芦霄鹏[20]将注意力机制和最小门控单元 MGU 引入 CNN, 进行文本情感分析。Tang 等[21]提出了一种联合方面类别的主题模型, 并使用最大熵来提高方面类别情感极性分类的表现。Tang 等[22]通过引入方面类别的词向量, 并改进传统的 LSTM 模型, 使模型达到在不使用情感词典情况下的最优结果。Xue 和 Li [23]提出了一种有效的神经单元, 通过将 CNN 和 GRU 结合使用, 对原始文本词向量和带有方面类别的词向量同时进行处理, 提高方面类别情感分类准确率。

### 2.2. 中文酒店在线评论情感分析

对中文酒店在线评论进行情感分析的研究中, Yao [24]提出一种基于 CNN-LSTM 的模型对中文酒店评论进行情感分析。马桂真和彭霞[25]通过运用 LDA 主题模型与构建情感词典相结合的方法, 对酒店评论数据进行游客关注的主题和情感倾向之间进行交叉挖掘分析。王宏鹏[26]分别提出构建酒店领域的情感词典和基于机器学习两种方法对酒店评论进行细粒度情感分析。池毛毛等[27]针对两大酒店预订平台上评论主题、主题社会网络 and 用户情感倾向的差异研究, 提出基于 LDA 的主题社会网络和情感分析方法。

## 3. 模型

本文提出的基于门控机制和 BiGRU 的 GCGAT 模型由 4 层网络结构构成, 分别为: 特征提取层、语义信息提取层、特征信息权重分配层和全连接层。模型以 CNN 和 BiGRU 为主体, 网络结构如图 1 所示。

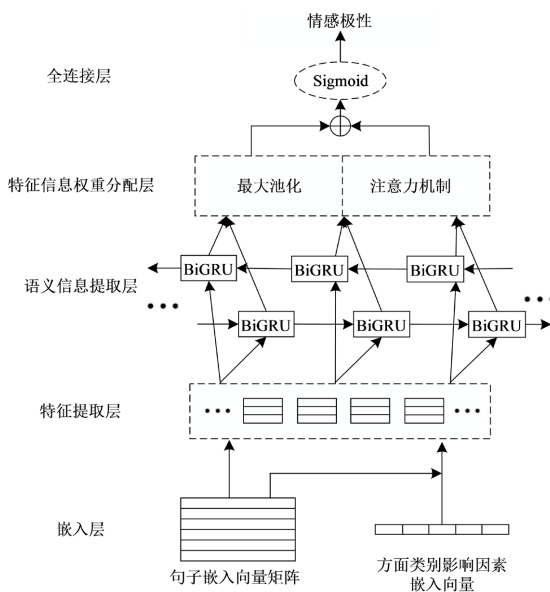


Figure 1. GCGAT model network structure

图 1. GCGAT 模型网络结构

### 3.1. 特征提取层

特征提取层为双通道结构, 其中, 通道 1 对句子嵌入矩阵进行特征提取, 通道 2 对同时嵌入方面类别影响因素词向量的嵌入矩阵进行特征提取。通过卷积层进行特征提取, 并使用门控机制控制输出, 门控机制选择 ReLU 和 GTU (Gated Tanh Unit) [28]两种, 其中 GTU 是基于门控机制所提出来的激活函数, 将 tanh 激活单元与 sigmoid 激活单元相结合, 不仅加快收敛速度, 还提高模型性能, 其计算公式如公式(1)。

$$h_i(X) = \tanh(X * W + b) \otimes \delta(X * V + c) \quad (1)$$

根据激活函数的函数特性, 通过 GTU 控制通道 1 的输出, 得到情感特征, 这样选择是因为细粒度酒店评论数据集中的情感倾向即包含负向情感倾向又包含正向情感倾向, 而 GTU 激活函数的值域为  $(-1, 1)$  可以避免情感倾向的丢失, 而对于方面类别, 评论中只有包含和不包含两种情况, 不包含即为 0, 包含即为大于 0, 这与 ReLU 激活函数的值域是相对应的, 所以将使用 ReLU 控制通道 2 的输出, 得到方面特征, 之后将二者输出进行相乘得到既包含情感特征又包含方面特征的特征信息。其结构如图 2 所示。

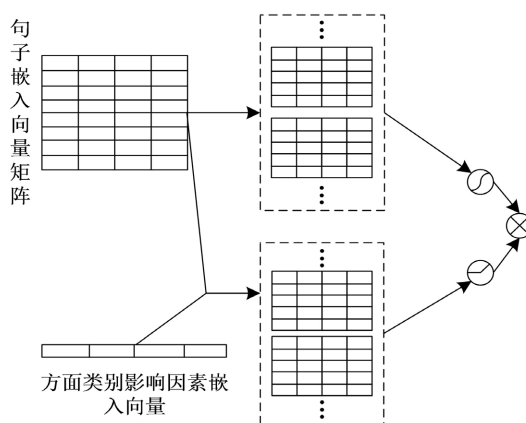


Figure 2. Feature extraction layer structure diagram

图 2. 特征提取层结构图

### 3.2. 语义信息提取层

语义信息提取层由 BiGRU 构成, 用来学习情感特征和方面类别特征的上下文信息, 其具有两层 GRU, 一层从顺序数据学习信息, 另外一层从反向输入数据学习信息。相较于传统循环神经网络(Recurrent Neural Networks, RNN)和长短时记忆网络(Long Short Term Memory Network, LSTM), GRU 引入了复位门和更新门来控制其输入和输出, 时刻结构如图 3 所示。

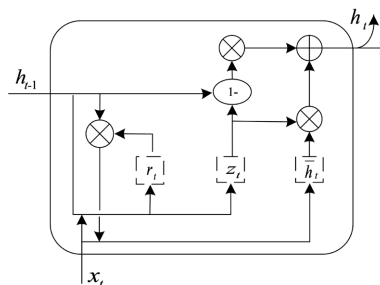


Figure 3. GRU structure diagram

图 3. GRU 结构图

其中,  $x_t$ 、 $h_{t-1}$  表示  $t$  时刻的输入,  $h_t$ 、 $y_t$  表示  $t$  时刻的输出, 参数更新公式见公式(2)、(3)、(4)、(5)、(6)。

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (2)$$

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (3)$$

$$\bar{h}_t = \tanh(W_h \cdot [r_t \times h_{t-1}, x_t]) \quad (4)$$

$$h_t = (1 - z_t) \times h_{t-1} + z_t \times \bar{h}_t \quad (5)$$

$$y_t = \sigma(W_o \cdot h_t) \quad (6)$$

其中,  $W_r$ ,  $W_z$ ,  $W_h$  和  $W_o$  是 GRU 的参数,  $[x, y]$  表示将  $x$  和  $y$  两个向量进行串联,  $r_t$  表示复位门,  $z_t$  表示更新门,  $\bar{h}_t$  表示候选隐藏层,  $\cdot$  表示元素乘法,  $\sigma$  表示 sigmoid 函数,  $\tanh$  表示 tanh 函数。

### 3.3. 特征信息权重分配层

特征信息权重分配层由最大池化和注意力机制组成, 通过最大池化获得最关键的特征信息, 注意力机制分配不同特征信息对方面类别情感倾向的影响权重, 以此保证不漏掉关键信息。注意力机制运用交互式建模可以通过判定对方面类别的情感做出贡献多少来赋予合适权重, 从而提高分类模型准确率, 其计算方法如公式(7)。

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}(\mathbf{Q}\mathbf{K}^T)\mathbf{V} \quad (7)$$

其中,  $\mathbf{Q}$  表示注意力机制的输入向量,  $\mathbf{K}$  表示文本中的每个键向量, 通过进行  $\mathbf{Q}$  和  $\mathbf{K}$  的相关性计算和归一化处理, 得到  $\mathbf{V}$  的注意力权重。

### 3.4. 全连接层

由于本文提出的方面类别情感分析任务是四分类任务, 所以选择具有 sigmoid 输出的全连接层进行最终的情感极性判定, 同时将特征信息权重分配层的输出结果进行拼接作为其输入。情感极性预测值  $\hat{y}_i$  的计算公式如公式(8)。

$$\hat{y}_i = \frac{1}{1 + e^{-(w_i^T C + b_i)}} \quad (8)$$

其中,  $\hat{y}_i$  表示第  $i$  个样本的预测值,  $w_i$  和  $b_i$  分别表示权重和偏置。

## 4. 数据

### 4.1. 数据获取

本文运用“八爪鱼”采集器这一数据采集软件从多个酒店预订平台获取酒店评论数据, 酒店的价格区间设定为 150~300 元之间, 酒店的范围包括北京、上海和深圳等一线城市。原始数据获取后通过数据清洗、数据去重、去除无效数据、以及去除数据长度大于 300 的数据后, 最终获得一万四千条有效数据。

### 4.2. 数据集建立

将最终获得的酒店评论数据进行分词和去停用词, 分词选择 jieba 分词方法, 停用词表选择哈工大停用词表、中文停用词表和百度停用词表的结合停用词表。将去停用词后的分词结果构建为词典, 用于 LDA-Jaccard 模型训练, 在 1~20 个主题个数之间选择最佳主题个数。训练结果如图 4 所示。

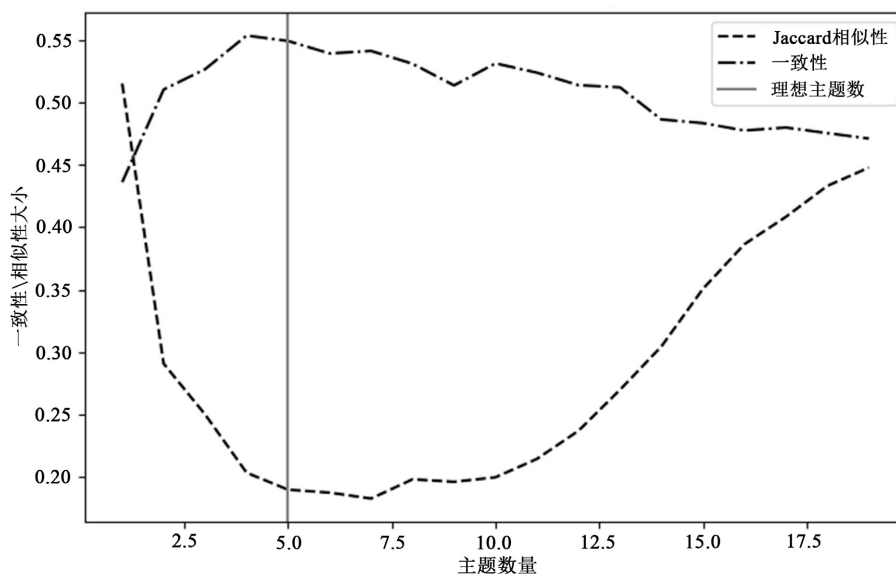


Figure 4. Topic-consistency, similarity change map

图 4. 主题 - 一致性、相似性变化图

根据图 4, 一致性得分会随着主题数的增加先升高后趋于平稳, 这是因为当主题数达到一定时, 主题内部一致性会达到最高, 再进行主题划分对一致性得分并不会产生太大影响, 但会使主题之间产生交叉, 为避免这种情况的出现, 本文引入了 Jaccard 相似性的计算, 通过计算主题之间的 Jaccard 相似性, 来判断主题之间的交叉程度。当主题数量为 5 时, 一致性得分与 Jaccard 相似性值之间的差值达到最大, 并且此时的一致性得分趋于平缓, Jaccard 相似性值接近最低, 说明当主题数为 5 时, 各主题内部的一致性达到稳定变化, 主题之间的交叉程度接近最低, 所以选取 5 个主题是合理的。此时, 各主题下频率最高的 25 个词如表 1 所示。

Table 1. High frequency words for each topic

表 1. 各主题下的高频词

主题	高频词
1	房间, 说, 晚上, 住, 隔音, 退房, 里, 真的, 太, 空调, 前台, 不好, 房, 差, 点, 卫生间, 好好, 找, 打扫, 窗户, 味道, 一点, 换, 阿姨, 吵
2	酒店, 前台, 入住, 体验, 感觉, 这家, 小姐姐, 装修, 选择, 房间, 喜欢, 住, 真的, 送, 贴心, 机器人, 帮, 帮忙, 朋友, 特别, 工作人员, 智能, 风格, 价格, 一家,
3	位置, 早餐, 西湖, 酒店, 地理位置, 周边, 吃, 很近, 房间, 出行, 北站, 旁边, 杭州, 步行街, 步行, 距离, 吃饭, 分钟, 火车站, 走, 出门, 楼下, 10, 住, 地方
4	服务, 房间, 免费, 舒服, 住, 满意, 升级, 很大, 接送, 便利, 新, 态度, 前台, 床, 停车, 交通, 酒店, 出差, 干净, 停车场, 深圳, 棒, 超级, 很棒, 地铁站
5	不错, 干净, 服务, 卫生, 环境, 房间, 推荐, 挺, 前台, 性价比, 服务态度, 下次, 热情, 设施, 特别, 高, 交通, 舒适, 入住, 值得, 齐全, 整洁, 整体, 安静, 还会

根据表 1, 经过综合考虑将 5 个方面类别名称分别定为环境好、服务周到、位置优越、交通便利和干净整洁, 并确定每个方面类别的影响因素如表 2 所示。

**Table 2.** Various categories of influencing factors  
**表 2.** 各方面类别影响因素

方面类别	影响因素
环境好	闹中取静, 设施, 环境, 装修, 智能家居
服务周到	服务周到, 前台, 接送, 班车, 大巴
位置优越	位置, 景点, 商场, 车站, 吃饭
交通便利	交通, 便利, 地铁站, 公交站, 停车场
干净整洁	干净, 卫生, 整洁

在数据标注过程中, 每一个方面类别都包含 4 种标签, 其中, -2 代表不涉及这一方面类别, -1 代表消极情感倾向, 0 代表中性情感倾向, 1 代表积极情感倾向。

## 5. 实验与分析

### 5.1. 参数设置与评价指标

本文选择 word2vec 模型训练词向量, 模型参数设置为: 词向量维数为 300, 训练算法选择 skip-gram 算法, 窗口大小为 15, 训练轮数为 20, 最小词频为 5。将训练好的词向量作为 GCGAT 模型的输入进行训练, 由于有 5 个方面类别, 所以本文选择将每个方面类别训练一次, 最终得到 5 个训练好的模型和结果。评价指标选择 Acc 值和 F1 值, 以及 ROC 曲线和 auc 值。Acc 值和 F1 值的计算公式为公式(9)、(10)。

$$Acc = \frac{TP + TN}{TP + FP + TN + FN} \quad (9)$$

$$F1 = \frac{2P * R}{P + R} \quad (10)$$

其中, 公式(9)中的指标分别为:  $TP$  表示真正例, 即预测为正例实际也为正例;  $TN$  表示真负例, 即预测为负例实际也为负例;  $FP$  表示假正例, 即预测为正例而实际为负例;  $FN$  表示假负例, 即预测为负例而实际为正例。公式(10)中的  $P$  表示精确度,  $R$  表示召回率, 二者计算公式为公式(11)、(12)。

$$P = \frac{TP}{TP + FP} \quad (11)$$

$$R = \frac{TP}{TP + FN} \quad (12)$$

auc 值定义为 ROC 曲线下与坐标轴围成的面积, ROC 曲线的横坐标为假阳率, 即真负例中判为正例的概率, 纵坐标为真阳率, 即真正例中判为正例的概率, 与召回率相同。

模型优化器选择 Adam 优化器, 学习率设置为 0.0001, 为避免过拟合, 模型选择 Spatial Dropout 1D [29] 和 Dropout, 二者值均设置为 0.2, Spatial Dropout 1D 用于对包含情感特征和方面特征的特征信息进行处理, Dropout 用于对包含情感特征、方面特征和上下文特征的特征信息进行处理。由于模型是多分类模型, 所以损失函数选择分类交叉熵, 计算公式为公式(13)。

$$loss = -\sum_i \sum_j y_i^j \log \hat{y}_i^j \quad (13)$$

其中,  $i$  表示句子编码,  $j$  表示分类编码,  $y$  表示目标分类,  $\hat{y}$  表示预测分类。

### 5.2. 对比实验

本文选择对比的模型包括: SVM, CNN, GRU, BiGRU, BiLSTM-aspect, 和 GCAE [9], 其中-aspect

表示模型输入的词向量是包含方面类别影响因素的词向量, 模型对比实验结果的准确率见表 3, F1 值见表 4。

**Table 3.** Comparative experimental accuracy

**表 3.** 对比实验准确率

model	环境好	服务周到	位置优越	交通便利	干净整洁
SVM	0.7999	0.9405	0.8234	0.9119	0.9582
CNN	0.8365	0.9108	0.8524	0.8887	0.8861
GRU	0.8524	0.8897	0.8897	0.8987	0.8921
BiGRU	0.8752	0.9526	0.8908	0.9329	0.9520
BiLSTM-aspect	0.8783	0.9602	0.8811	0.9381	0.9651
GCAE	0.8772	0.9586	0.8984	0.9388	0.9696
GCGAT	<b>0.9049</b>	<b>0.9661</b>	<b>0.9160</b>	<b>0.9554</b>	<b>0.9796</b>

**Table 4.** The F1 values of the experiment were compared

**表 4.** 对比实验 F1 值

model	环境好	服务周到	位置优越	交通便利	干净整洁
SVM	0.8032	0.9417	0.8280	0.9141	0.9595
CNN	0.8214	0.9088	0.8566	0.8861	0.8899
GRU	0.8308	0.8816	0.8427	0.8811	0.8792
BiGRU	0.8723	0.9513	0.8889	0.9327	0.9511
BiLSTM-aspect	0.8691	0.9586	0.8767	0.9349	0.9636
GCAE	0.8811	0.9596	0.9018	0.9423	0.9710
GCGAT	<b>0.8995</b>	<b>0.9647</b>	<b>0.9124</b>	<b>0.9527</b>	<b>0.9753</b>

由表 3 和表 4 可以发现, SVM 作为机器学习中的一种经典算法, 在服务周到、交通便利和干净整洁这三个方面类别中准确率和 F1 值甚至比 CNN 和 GRU 的更优, 并且准确率和 F1 值在服务周到这一方面类别上分别为 0.9405 和 0.9417, 在交通便利这一方面类别上分别为 0.9119 和 0.9141, 在干净整洁这一方面类别上二者均取得最优结果, 分别为 0.9582 和 0.9595。CNN 的准确率和 F1 值在服务周到这一方面类别上结果更好, 分别为 0.9108 和 0.9088, 计算得出比 GRU 的准确率和 F1 值分别提升了 2.1%和 2.7%。BiGRU 的准确率比 GRU 平均提升了 3.6%, F1 值比 CNN 平均提升了 4.7%, 并且在服务周到这一方面类别上取得了最优结果, 准确率和 F1 值分别为 0.9526 和 0.9513, 相较于 SVM, 分别提升了 1.2%和 1%。BiLSTM-aspect 基于 BiLSTM 引入了方面影响因素词向量, 使模型分类结果进一步提升, 其准确率相较于 SVM 平均提升了 3.8%, F1 值相较于 CNN 平均提升了 4.8%, 在服务周到和干净整洁这两个方面类别上准确率和 F1 值均达到了 0.95 以上。GCAE 模型通过门控机制控制方面信息和情感信息的获取, 极大提升了 CNN 的准确率和 F1 值, 结果分别平均提升了 5.4%和 5.9%, 在位置优越这一方面类别上 F1 值比 BiLSTM-aspect 的 F1 值提升了 2.5%。同时, 可以发现各模型在此数据集上的准确率和 F1 值都达到了 0.8 左右, 甚至超过 0.9, 这说明该数据集的有效性。再将本文提出的深度学习模型 GCGAT 模型与这些模型的结果进行对比发现, GCGAT 模型的准确率和 F1 值均有所提升, 通过计算得出, 比 GRU 的准确率平



均提升了 6%，比 BiGRU 的准确率平均提升了 2.4%，比 GCAE 的准确率平均提升了 1.6%，并且在各方面类别情感分类中的准确率都达到了 0.9 以上，在干净整洁这一类别中最高，达到了 0.9796。GCGAT 模型的 F1 值比 SVM 的 F1 值平均提升了 5.2%，比 CNN 的 F1 值平均提升了 6.8%，比 BiLSTM-aspect 的 F1 值平均提升了 2%，在服务周到、交通便利和干净整洁这三个方面类别上均达到了 0.95 以上，虽然在环境好这一方面类别上的结果最低，但是仍达到了 0.8995，证明了 GCGAT 模型的有效性。

进一步可以发现，模型在服务周到和干净整洁这两个方面类别上面的准确率和 F1 值比其他方面类别上面的准确率和 F1 值会更高，这可能是因为在标注服务周到和干净整洁这两个方面类别情感倾向时，评论中这两个方面的影响因素表示明确，而其他的方面类别情感倾向在标注过程中可能会因为评论的模糊而存在一些混淆，是模型在进行分析时产生偏差。

接下来选择 SVM, BiLSTM-aspect 和 GCGAT 作为对比模型与 GCGAT 模型比较 ROC 曲线和 auc 值，这样选择是以 SVM 代表机器学习，BiLSTM-aspect 和 GCGAT 代表深度学习并且嵌入方面类别词向量。由于本文为多分类任务，所以在绘制 ROC 曲线和计算 auc 值时选择“micro”模式，并且将五个方面类别分类情况放在一起考虑，得到整体的假阳率和真阳率。由图 5 可以看出 GCGAT 模型取得了更好的效果，auc 值达到了 0.99，说明模型可以有效的将情感极性判定正确。BiLSTM-aspect 模型也取得了很好的效果，auc 值为 0.97，说明考虑上下文信息很重要。

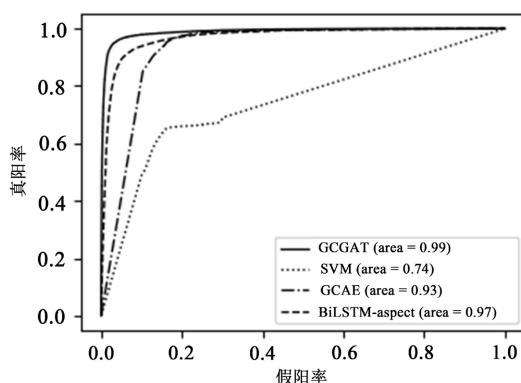


Figure 5. ROC curve comparison diagram

图 5. ROC 曲线对比图

### 5.3. 消融实验

为了检验 GCGAT 模型中引入的方面类别影响因素词向量、GTU 激活函数、BiGRU 和注意力机制的有效性，接下来对 GCGAT 模型进行了消融实验。其中，GCGAT/GTU 模型是指将 GCGAT 模型中的 GTU 激活函数替换为 tanh 激活函数；GCGAT/BiGRU 模型是指将 GCGAT 模型中的 BiGRU 去除；GCGAT/aspect 模型是指将 GCGAT 模型中的方面类别影响因素的词向量去除；GCGAT/attention 模型是指将 GCGAT 模型中的 attention 机制去除，仅通过最大池化保留关键的信息。实验结果的准确率见表 5，F1 值见表 6。

Table 5. Accuracy of ablation experiment

表 5. 消融实验准确率

model	环境好	服务周到	位置优越	交通便利	干净整洁
GCGAT/GTU	0.9021	0.9573	0.9129	0.9495	0.9789
GCGAT/BiGRU	0.9029	0.9633	0.9143	0.9488	0.9775

Continued

GCGAT/aspect	0.8835	0.9571	0.9004	0.9440	0.9640
GCGAT/attention	0.8821	0.9567	0.8997	0.9443	0.9695
GCGAT	<b>0.9049</b>	<b>0.9661</b>	<b>0.9160</b>	<b>0.9551</b>	<b>0.9796</b>

Table 6. F1 of ablation experiment

表 6. 消融实验 F1 值

model	环境好	服务周到	位置优越	交通便利	干净整洁
GCGAT/GTU	0.8595	0.9386	0.8754	0.9314	0.9462
GCGAT/BiGRU	0.8788	0.9615	0.9013	0.9466	0.9743
GCGAT/aspect	0.876	0.9535	0.8867	0.9402	0.9588
GCGAT/attention	0.8781	0.9502	0.901	0.9392	0.9678
GCGAT	<b>0.8995</b>	<b>0.9647</b>	<b>0.9124</b>	<b>0.9527</b>	<b>0.9753</b>

通过表 5 和表 6 可以得出, 方面类别影响因素向量和注意力机制的引入会对准确率的提升产生更大的影响, 而 GTU 激活函数的引入会对 F1 值的提升产生更大影响, 同时 BiGRU 的引入在环境好和位置优越这两个方面类别上对 F1 值的提高也起到了巨大作用。根据计算得到, GCGAT 模型的准确率比 GCGAT/aspect 模型的准确率平均提高了 1.5%, 比 GCGAT/attention 模型的准确率平均提高了 1.4%, 并且均在环境好这一方面类别上提升最多, 比 GCGAT/aspect 模型提升了 2.1%, GCGAT/attention 模型提升了 2.3%。GCGAT 模型的 F1 值比 GCGAT/GTU 模型的 F1 值平均提高了 3.1%, 并且也在环境好这个方面类别中提升最多, 提升了 4%, 同时在位置优越这一方面类别上提升了 3.7%。通过与 GCGAT/BiGRU 模型进行比较, GCGAT 模型在环境好和位置优越这两个方面类别情感分析中的 F1 值分别提升了 2%和 1%。

根据上述分析可以总结出, 由于此时模型提取了方面类别特征和与方面类别特征相关的情感特征, 考虑了评论中的上下文信息, 最后保留最重要的特征信息以及其他关键信息, 并利用这些信息进行情感极性判定, 这样做使方面类别与其情感倾向更加匹配, 极大地提升了模型性能。

## 6. 结论

本文从方面级情感分析中的方面类别情感分析出发, 提出一个同时融合方面类别信息, 情感信息和上下文信息的深度学习模型 GCGAT 模型, 并通过该模型对中文酒店评论进行方面类别情感分析。该模型有四层网络结构, 分别为: 特征提取层、语义信息提取层、特征信息权重分配层和全连接层, 主体由 CNN 和 BiGRU 构成, 通过并列使用两个卷积层同时提取评论中的情感特征和方面类别特征, 将两种特征合并后, 通过 BiGRU 学习合并特征在评论中的上下文的信息, 运用最大池化保留的最重要的特征信息和注意力机制保留的其他关键信息, 最后将全部信息输入到全连接层中进行方面类别的情感极性判定。在数据集方面, 本文运用 LDA-Jaccard 模型对经过预处理的数据进行分类并确定方面类别, 在此基础上进行数据标注。通过与其他模型进行对比, 可以发现 GCGAT 模型取得了更好的结果, 准确率比 BiGRU 的准确率平均提升了 2.4%, 同时, 在全部方面类别上均达到了 0.9 以上。F1 值比 BiLSTM-aspect 的 F1 值平均提升了 2%, 在服务周到、交通便利和干净整洁这三个方面类别上均达到了 0.95 以上。并且 GCGAT 模型在干净整洁这一方面类别上得到了最优结果, 准确率达到 0.9796, F1 值达到了 0.9753。目前预训

练模型在自然语言处理任务中取得了广泛并有效的使用, 所以, 在未来工作中将研究将预训练模型引入到酒店评论方面及情感分析当中。

## 参考文献

- [1] Wang, J., Du, J.P., Shao, Y.X. and Li, A. (2022) Sentiment Analysis of Online Travel Reviews Based on Capsule Network and Sentiment Lexicon. In: Zhang, L., Yu, W., Jiang, H. and Laili, Y., Eds., *CINT 2022: Intelligent Networked Things*, Springer, Singapore, 92-106.
- [2] 邓慈云, 余国清. 基于朴素贝叶斯的影评情感分析研究[J]. 智能计算机与应用, 2023, 13(2): 210-212, 217.
- [3] Tong, X., Li, Y.X., Li, J.Y., et al. (2022) What Are People Talking about in #BlackLivesMatter and #StopAsianHate? Exploring and Categorizing Twitter Topics Emerging in Online Social Movements through the Latent Dirichlet Allocation Model. *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, Oxford, 19-21 May 2021, 723-738. <https://doi.org/10.1145/3514094.3534202>
- [4] 孔繁钰, 陈纲. 基于改进双向 LSTM 的评教文本情感分析[J]. 计算机工程与设计, 2022, 43(12): 3580-3587.
- [5] Wei, J.Y., Liao, J., Yang, Z.F., et al. (2020) BiLSTM with Multi-Polarity Orthogonal Attention for Implicit Sentiment Analysis. *Neurocomputing*, 383, 165-173. <https://doi.org/10.1016/j.neucom.2019.11.054>
- [6] 刘继, 顾凤云. 基于 BERT 与 BiLSTM 混合方法的网络舆情非平衡文本情感分析[J]. 情报杂志, 2022, 41(4): 104-110.
- [7] Phan, H.T., Nguyen, N.T. and Hwang, D. (2022) Aspect-Level Sentiment Analysis Using CNN over BERT-GCN. *IEEE Access*, 10, 110402-110409. <https://doi.org/10.1109/ACCESS.2022.3214233>
- [8] 曹卫东, 潘红坤. 采用稀疏自注意力机制和 BiLSTM 模型的细粒度情感分析[J]. 计算机应用与软件, 2022, 39(12): 187-194.
- [9] Yang, J. and Yang, J. (2020) Aspect Based Sentiment Analysis with Self-Attention and Gated Convolutional Networks. *2020 IEEE 11th International Conference on Software Engineering and Service Science*, Beijing, 16-18 October 2020, 146-149. <https://doi.org/10.1109/ICSESS49938.2020.9237640>
- [10] Fei, H., Li, F., Li, C.L., et al. (2022) Inheriting the Wisdom of Predecessors: A Multiplex Cascade Framework for Unified Aspect-Based Sentiment Analysis. *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence Main Track*, Vienna, 23-29 July 2022, 4121-4128. <https://doi.org/10.24963/ijcai.2022/572>
- [11] 中华人民共和国国家统计局. 中国统计年鉴[M]. 北京: 中国统计出版社, 2022.
- [12] 邬小燕. 基于 Stacking 的酒店评论情感分析研究[D]: [硕士学位论文]. 重庆: 重庆大学, 2019.
- [13] 高华玲, 张晶. 基于情感词典的酒店评论情感分析与可视化[J]. 软件, 2021, 42(1): 45-47, 66.
- [14] 丁美荣, 冯伟森, 黄荣翔, 等. 基于预训练模型和基础词典扩展的酒店评论情感分析[J]. 计算机系统应用, 2022, 31(11): 296-308.
- [15] Xue, W. and Li, T. (2018) Aspect-Based Sentiment Analysis with Gated Convolutional Networks. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, Melbourne, July 2018, 2514-2523. <https://doi.org/10.18653/v1/P18-1234>
- [16] Vaswani, A., Shazeer, N., Parmar, N., et al. (2017) Attention Is All You Need. arXiv: 1706.03762.
- [17] Kiritchenko, S., Zhu, X.D., Cherry, C. and Mohammad, S. (2014) NRC-Canada-2014: Detecting Aspects and Sentiment in Customer Reviews. *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, Dublin, August 2014, 437-442. <https://doi.org/10.3115/v1/S14-2076>
- [18] 王新宇. 基于情感词典与机器学习的旅游网络评价情感分析研究[J]. 计算机与数字工程, 2016, 44(4): 578-582, 766.
- [19] Mikolov, T., Sutskever, I., Chen, K., et al. (2013) Distributed Representations of Words and Phrases and Their Compositionality. arXiv: 1310.4546.
- [20] 徐菲菲, 芦霄鹏. 结合卷积神经网络和最小门控单元注意力的文本情感分析[J]. 计算机应用与软件, 2020, 37(9): 75-80, 125.
- [21] Tang, F.L., Fu, L.Y., Yao, B., et al. (2019) Aspect Based Fine-Grained Sentiment Analysis for Online Reviews. *Information Sciences*, 488, 190-204. <https://doi.org/10.1016/j.ins.2019.02.064>
- [22] Tang, D.Y., Qin, B., Feng, X.C., et al. (2016) Effective LSTMs for Target-Dependent Sentiment Classification. *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, Osaka, December 2016, 3298-3307.

- 
- [23] Xue, W. and Li, T. (2018) Aspect-Based Sentiment Analysis with Gated Convolutional Networks. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, Melbourne, July 2018, 2514-2523. <https://doi.org/10.18653/v1/P18-1234>
- [24] Yao, L.S. (2022) Sentiment Analysis Based on CNN-LSTM Hotel Reviews. *Journal of Physics: Conference Series*, **2330**, Article ID: 012018. <https://doi.org/10.1088/1742-6596/2330/1/012018>
- [25] 马桂真, 彭霞. 基于酒店评论大数据的游客评论主题挖掘与情感分析——以北京五星级酒店为例[J]. 北京联合大学学报, 2021, 35(2): 58-68.
- [26] 王宏鹏. 基于词典与机器学习的酒店评论情感分析的研究[D]: [硕士学位论文]. 大连: 大连交通大学, 2020.
- [27] 池毛毛, 潘美钰, 王伟军. 共享住宿与酒店用户评论文本的跨平台比较研究: 基于 LDA 的主题社会网络和情感分析[J]. 图书情报工作, 2021, 65(2): 107-116.
- [28] Dauphin, Y.N., Fan, A., Auli, M. and Grangier, D. (2017) Language Modeling with Gated Convolutional Networks. *Proceedings of the 34th International Conference on Machine Learning*, **70**, 933-941.
- [29] Jonathan, T., Ross, G., Arjun, J., *et al.* (2015) Efficient Object Localization Using Convolutional Networks. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, 7-12 June 2015, 648-656.