



开放科学  
(资源服务)  
标识码  
(OSID)

# 酒店在线评论在 TextCNN 和 SVM 方法下的情感分类研究

吴宋体<sup>1</sup> 陈家虎<sup>2</sup>

1. 桂林学院工商管理学院 桂林 541004;
2. 广州中广计算机科技有限公司 广州 510000

**摘要:** [目的/意义] 在线评论中蕴含丰富的用户情感信息, 对酒店行业服务优化与管理决策具有重要意义。本文探讨了 TextCNN 与 SVM 两种情感分类方法在中文酒店评论中的应用效果。[方法/过程] 基于携程平台的真实酒店评论数据, 经过数据预处理与词向量化处理, 分别构建并训练了 TextCNN 和 SVM 情感分类模型, 通过准确率、精确率、召回率和 F1 值等指标评估模型性能, 并分析了错误分类样本。[局限] 训练数据量较小, 可能影响模型的泛化能力。[结果/结论] 实验结果表明, TextCNN 和 SVM 在测试数据中分别达到了 96% 和 93% 的准确率, 其中 TextCNN 在验证集中的准确率 (86%) 优于 SVM (83%)。错误样本分析表明, 两种模型在处理复杂或中性表达情感时存在误判, TextCNN 对复杂语境的理解优于 SVM。总体而言, 研究验证了深度学习与传统机器学习在酒店评论情感分类中的有效性与差异, 为酒店行业情感分析模型选择提供了实证依据。

**关键词:** 情感分析; TextCNN; 在线评论; SVM; 酒店服务

**中图分类号:** G35; TP391

## Research on Sentiment Classification of Hotel Online Reviews Using TextCNN and SVM Methods

WU Songti<sup>1</sup> CHEN Jiahu<sup>2</sup>

1. School of Business Administration, Guilin University, Guilin 541004, China;
2. GuangZhou ZhongGuang Computer Technology Co., Ltd, Guangzhou 510000, China

**Abstract:** [Objective/Significance] Online reviews contain rich user sentiment information, which is of great significance for service optimization and management decision-making in the hotel industry. This paper explores the application effects of two sentiment classification methods, TextCNN and SVM, in Chinese hotel reviews. [Methods/Processes] Based on real hotel review data from the Ctrip platform, TextCNN and SVM sentiment classification models were constructed and trained after data

**作者简介** 吴宋体 (1988-), 通讯作者, 硕士, 工程师, 主要研究方向为商务大数据与运营管理, E-mail: wusongti@163.com; 陈家虎 (1982-), 硕士, 主要研究方向为企业信息化。

**引用格式** 吴宋体, 陈家虎. 酒店在线评论在 TextCNN 和 SVM 方法下的情感分类研究 [J]. 情报工程, 2026, 12(1): 59-70.

preprocessing and word vectorization. Model performance was evaluated through metrics such as accuracy, precision, recall, and F1-score, and misclassified samples were analyzed. [Limitation] The small volume of training data may affect the generalization capability of the models. [Results/Conclusions] Experimental results demonstrate that TextCNN and SVM achieved accuracies of 96% and 93% on the test data, respectively, with TextCNN outperforming SVM on the validation set (86% compared to 83%). Misclassified sample analysis indicates that both models exhibit misjudgments when processing complex or neutral expressions of sentiment, and TextCNN demonstrates superior understanding of complex contexts compared to SVM. Overall, the study validates the effectiveness and differences between deep learning and traditional machine learning in the sentiment classification of hotel reviews, providing an empirical basis for model selection in hotel industry sentiment analysis.

**Keywords:** Sentiment Analysis; TextCNN; Online Reviews; SVM; Hotel Services

## 引言

随着互联网和社交媒体的普及，在线评论已经成为用户表达情感和分享体验的重要平台，特别是在酒店行业，用户的评论对酒店服务优化和管理决策具有重要参考价值。酒店评论不仅直接影响消费者的决策行为，也为酒店管理者提供了宝贵的反馈信息<sup>[1-2]</sup>。如何高效地从海量在线评论文本中挖掘用户情感信息，已成为服务行业（如酒店业）服务优化和管理决策的重要研究课题<sup>[3]</sup>。

情感分析作为一种文本挖掘技术，旨在从文本中提取用户的情绪倾向，已广泛应用于旅游、商业、政务等多个领域<sup>[1, 4-5]</sup>。传统的情感分析方法多基于情感词典进行规则匹配，这种方法依赖人工构建的情感词典，难以处理新兴词汇和复杂语义<sup>[6]</sup>。随着机器学习的发展，支持向量机（Support Vector Machine, SVM）等传统方法在情感分类中得到了广泛应用，其因分类精度高、计算效率优良而受到青睐<sup>[5, 7]</sup>。然而，这些机器学习方法通常依赖高质量的人工特征工程，存在一定的局限性<sup>[3, 5]</sup>。

近年来，深度学习特别是卷积神经网络（Convolutional Neural Network, CNN）的应用为情感分析带来了新的突破。文本卷积神经网络（Text Convolutional Neural Network, TextCNN）模型因其简单高效的结构和强大的语义特征提取能力，逐渐成为文本情感分析中的主流方法之一<sup>[8-9]</sup>。与传统的机器学习方法相比，TextCNN能够自动学习并捕捉文本中的上下文特征，显著减少了对人工特征工程的依赖，尤其适用于中文酒店评论这一语言表达灵活的场景<sup>[9-11]</sup>。

尽管已有大量研究针对情感分析方法在中文评论中的应用进行探索，但在酒店评论领域，基于深度学习与传统机器学习方法的系统性比较仍较为缺乏。尤其是在中文评论中，广泛存在语言积极性偏见（Language Positive Bias, LPB）和多维度情感表达等问题，影响了情感分类模型的泛化能力和应用效果<sup>[12-13]</sup>。因此，本文基于携程平台的真实酒店评论数据，分别构建TextCNN和SVM模型，对两者在情感分类中的表现进行对比分析，旨在为酒店行业提供适用的情感分析模型，并为未来相关研究提

供理论参考和实践指导。

## 1 数据准备

### 1.1 数据采集

本文以酒店在线评论为分析对象，所涉及的数据主要包括两类：

#### (1) 模型情感分类效果的验证数据

为评估情感分类模型的实际效果，选取携程在线预订平台上的酒店用户评论作为待分类文本数据。评论样本涵盖多家主流连锁酒店，包括汉庭、城市便捷、希尔顿和亚朵等。通过网页信息提取技术收集了截至 2025 年的大量评论文本，筛选出具有明显情感倾向的样本约 1500 条，其中评分为 1 分（代表“非常不满意”）的评论和评分为 5 分（代表“非常满意”）的评论分别约为 750 条。此类极端评分的评论往往情绪表达明确，情感极性显著，适用于情感倾向的分类验证任务。

#### (2) 模型训练所用语料库

模型训练所使用的基础语料为中科院计算

所谭松波整理发布的中文酒店评论情感语料库（ChnSentiCorp\_htl\_all）以及 GitHub 仓库上公开的酒店评论情感语料库，如表 1 所示。

表 1 模型训练语料库来源情况

数据集名称	来源	数量
ChnSentiCorp_htl_all	中科院计算所 谭松波整理	正负情感共约 1000条
NLP_data/携程酒店 评价.csv	GitHub仓库	正负情感共约 5000条
Lstm/hotel_discuss2. csv	GitHub仓库	正负情感共约 6600条

### 1.2 数据预处理

为确保情感分类模型的性能，在将原始文本输入模型之前，需对语料进行系统化的预处理。无论是用于训练深度学习模型的语料，还是用于最终分类评估的酒店评论数据，均依照统一流程完成清洗与规范化操作，如图 1 所示。该流程主要包括噪声剔除、词形还原、重复项过滤及中文分词等步骤，主要采用 Python 程序完成这些操作。

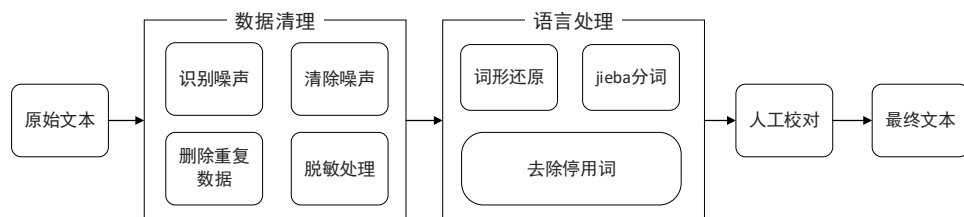


图 1 数据预处理流程

所有数据在经过 Python 程序的数据清洗和语言处理后，再由研究人员对结果逐条进行人工校对，包括但不限于情感分类的正确性与分词的合理性进行判断与校正，以保障分词精度与语义一致性。经此流程处理后，原始评论数

据转换为更加干净且结构化的文本格式。

表 2 为处理后的数据统计，清洗后数据总量虽略有减少，主要损耗来自去除噪声和重复记录，但文本质量和信息密度显著提高，为模型训练和测试提供了可靠的数据基础。

表 2 处理后的数据统计

数据集	情感类别	合计
模型训练语料库	正向5544条	11460条
	负向5916条	
酒店用户在线评论	正向700条	1400条
	负向700条	

通过上述数据采集与预处理，我们建立了涵盖训练和测试所需的高质量语料，为不同方法的对比实验创造了统一和公平的条件。

## 2 实验设计与评价指标

### 2.1 实验设计

为了深入探究不同方法在酒店在线评论情感分类任务中的应用效果，本研究设计了一套系统的对比实验。实验核心在于评估 TextCNN 与 SVM 的性能差异及各自优势。整体实验流程严格按照逻辑推进，如图 2 所示：首先，在前期预处理得到的高质量语料基础上，分别构建并训练 TextCNN 和 SVM 情感分类模型。接着利用准确率、精确率、召回率和 F1 值等综合指标对模型开展详尽的性能分析。本实验设计旨在通过统一和公平的测试条件，为酒店行业高效挖掘用户情感信息以及选择适用的情感分析模型提供科学的实证依据。

#### 2.1.1 构建TextCNN情感分类模型

TextCNN 是一种基于 CNN 的文本分类模型，最早由 Kim<sup>[9]</sup> 提出，其核心思想是通过多个不同尺寸的卷积核提取文本局部 n-gram 特征，并通过最大池化选取显著特征，再连接至全连接层进行情感极性分类，其模型结构如图 3 所示。

相比于传统特征工程方法，TextCNN 能够自动学习文本语义表示，具有较强的上下文感知能力和特征提取效率，尤其适用于中文在线评论这类语序灵活、表达多变的短文本分析<sup>[8, 10-11]</sup>。

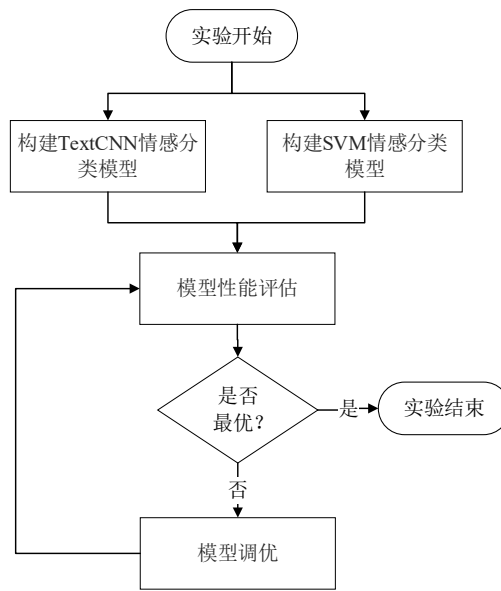


图 2 实验步骤

评论文本经过分词和词向量化处理后，输入 TextCNN 模型进行训练与预测，从而实现酒店评论正负情感的二分类判断。构建 TextCNN 情感分类模型的具体步骤如下：

#### (1) 参数设置

本文在构建模型时，充分参考了前人文献中关于 TextCNN 的改进实践<sup>[9-11, 14]</sup>，结合实验数据规模、评论文本长度分布和过拟合风险等因素，确定了 TextCNN 模型参数配置，如表 3 所示。

#### (2) 数据填充和截取

鉴于深度学习模型对输入数据维度的对齐要求，需将原始文本序列统一映射为固定长度的数值张量。在文本预处理阶段，本文对序列进行了长度标准化处理：对于超出设定长度的

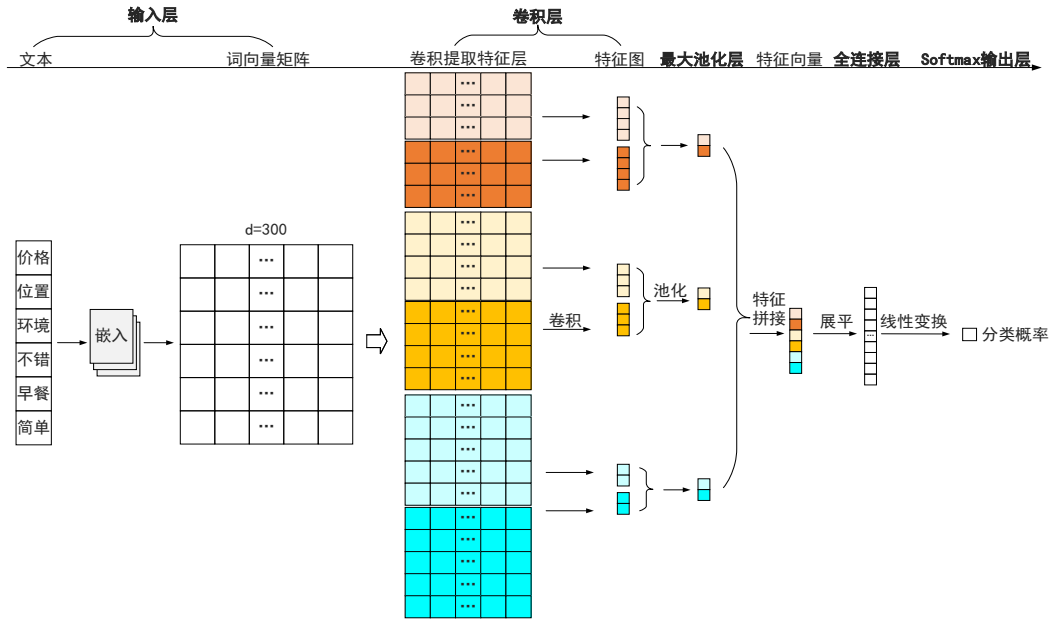


图3 TextCNN 情感分类模型

表3 模型参数设置说明

模块	参数项	设置值	设计说明
数据集划分	训练样本数量	11460	其中每一种数据集下正负向数据各占一半
	test_size	0.2	按照8:2的比例划分为训练集和测试集
	词向量维度	300	常见经验值，平衡表示能力与计算效率
嵌入层	最大文本长度	60	经过评估，可覆盖95%以上的文本序列
	是否可训练	是 (trainable=True)	嵌入层权重参与训练，更好适配具体语料
	卷积核数量	128	适配中等数据规模，避免过拟合
卷积层	卷积窗口大小	3, 4, 5	提取不同粒度的局部特征
	激活函数	ReLU	经典非线性激活函数
	填充方式	same	保持输入输出长度一致
	正则化方式	L2正则, $\lambda=0.01$	缓解过拟合风险
池化层	池化方式	GlobalMaxPooling1D	提取各卷积通道的最强激活，简化结构
特征融合层	融合方式	Concatenate	多尺度特征拼接以增强语义表达能力
Dropout层	比例	0.5	随机丢弃防止模型对局部特征过拟合
输出层	单元数	1 (Sigmoid)	二分类输出，适用于正负情感判别
	正则化方式	L2正则 $\lambda=0.01$	统一正则策略，抑制参数冗余
	批处理大小	batch_size=128	兼顾训练效率与显存消耗
训练设置	最大训练轮数	Epochs=20	初始设定，结合早停机制动态终止
	验证集划分	20%	用于监控模型泛化能力
学习策略	学习率调整	ReduceLROnPlateau (factor=0.1, patience=3)	动态调整学习率，提升收敛速度
	早停机制	EarlyStopping (patience=3)	防止过度训练导致性能退化

序列执行后方截断，保留核心语义信息；对于长度不足的序列则采用左侧补零方式进行填充。经统计，虽然语料库中原始文本序列最长达310个字符，但95%的样本长度分布在60个字

符以内。为在高维特征提取与计算效率之间取得平衡，本文将模型输入的最大序列长度统一设定为60。数据填充与截取的具体逻辑如图4所示。

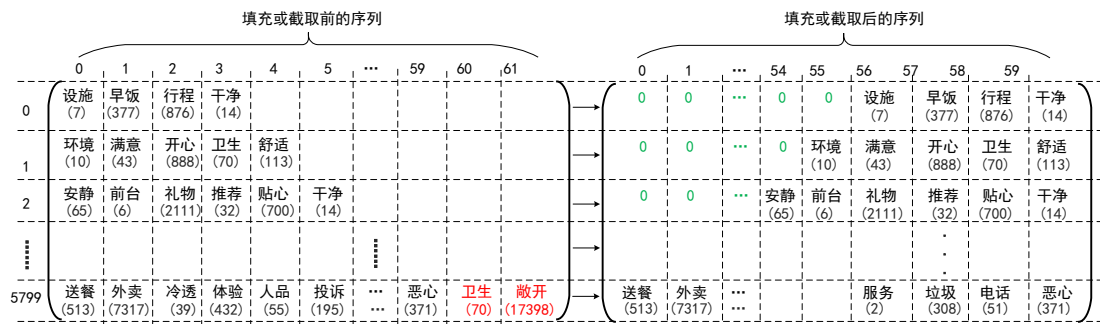


图4 数据填充或截取示意图

### (3) 模型训练

在模型训练阶段，首先将输入文本通过嵌入层转化为词向量矩阵，对于输入的词语序列  $[w_1, w_2, \dots, w_n]$ ，每个词通过嵌入层映射为一个  $d$  维词向量，嵌入矩阵可通过公式(1)表示为：

$$X=[e_{w_1}, e_{w_2}, \dots, e_{w_n}] \in R^{n \times d} \quad (1)$$

其中， $n$  为句子长度， $d=300$  为词向量维度， $e_{w_i}$  表示第  $i$  个词的嵌入向量。

然后分别使用不同大小的卷积核进行卷积操作，提取不同尺度的局部特征。给定窗口大小为  $h$ 、卷积核为  $W \in R^{h \times d}$ ，第  $i$  个位置的卷积特征通过公式(2)表示为：

$$c_i=f(w, X_{i:i+h-1})+b) \quad (2)$$

其中， $f$  为非线性激活函数（如 ReLU）， $(\cdot)$  表示加权和， $X_{i:i+h-1}$  表示从第  $i$  个词开始的  $h$  个词的嵌入子矩阵。

每个卷积核对应的特征图经过最大池化后，得到一组固定长度的特征向量。将各个池化后的特征向量拼接成一个统一的特征向量，经过

展平后输入到全连接层。随后，通过线性变换将特征向量映射为两个类别的得分（logits），最后使用 Softmax 函数将得分归一化为正面和负面的概率，取概率最大的类别作为最终情感分类结果。公式表示为：

$$\hat{y}=\sigma(w^T z+b) \quad (3)$$

其中， $z$  为拼接后的特征向量， $\sigma(\cdot)$  为 Sigmoid 激活函数， $\hat{y} \in (0,1)$ ，表示正向情感的概率。

### (4) 模型训练日志分析

训练共设置 20 个 epoch，并引入早停机制以防止模型过拟合，当验证集的性能在连续若干轮的训练中不再提升时则提前终止训练。表4显示，模型在前几轮训练中快速收敛：第1轮训练准确率为67%，验证准确率为69%；第3轮时，训练损失下降至39%，验证损失降至46%，验证准确率提升至85%。模型在第6至8轮左右达到最优性能，训练准确率达到98%，验证准确率稳定在84%左右，验证损失维持在46%，表明模型在本任务中的训练过程稳定、收敛迅速，并且泛化能力良好。

表 4 TextCNN 模型训练日志

轮次	训练损失	准确率	验证损失	验证准确率
01/20	1.59	0.67	0.69	0.77
02/20	0.53	0.85	0.49	0.83
03/20	0.39	0.91	0.46	0.85
04/20	0.32	0.94	0.45	0.86
05/20	0.28	0.96	0.45	0.85
06/20	0.25	0.97	0.46	0.84
07/20	0.23	0.98	0.46	0.85
08/20	0.21	0.98	0.46	0.84

### 2.1.2 构建SVM情感分类模型

SVM 是一种传统的监督学习算法，广泛应用于文本分类和情感分析任务，其基本原理是通过构建最优分类超平面以最大化类别间的间隔，从而提升分类器的泛化能力，如图 5 所示。SVM 在处理高维稀疏文本特征方面表现稳定，尤其适用于中小样本情感分类任务<sup>[3, 11]</sup>。

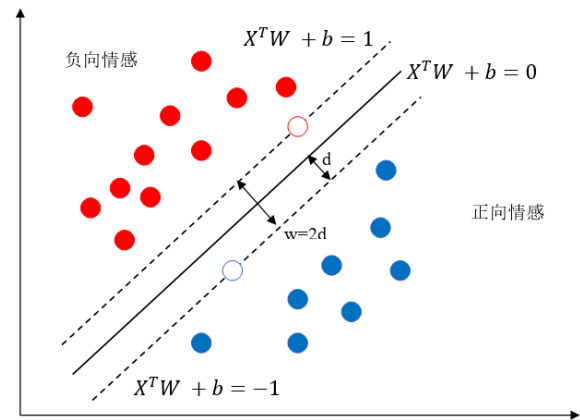


图 5 SVM 支持向量机原理

首先对酒店评论进行中文分词并提取 TF-IDF 或 Word2Vec 特征向量，然后输入 SVM 模型进行训练，通过核函数映射实现情感倾向的二分类。构建 SVM 情感分类模型的具体步骤如下。

#### (1) 参数设置

相关参数在参考其他研究人员设置的基础上，经过多轮参数调优后，得出本次构建 SVM 模型的最佳参数配置，如表 5 所示。

表 5 模型训练相关参数设置

模块	参数项	设置值	设计说明
数据集划分	训练样本数量	11460	为控制训练集正负样本平衡，每一种数据集下正负向数据各占一半
	test_size	0.2	按照 8:2 的比例划分为训练集和测试集
特征提取	max_features	5000	控制 TF-IDF 向量维度上限，限制模型计算复杂度
	ngram_range	(1, 2)	提取 1 元和 2 元词组特征，增强模型对上下文的表达能力
模型选择	分类器	SVC(probability=True)	使用支持向量机，适合稀疏高维文本特征，保留概率预测能力
	kernel	linear	使用线性核，适合大规模文本数据且训练速度更快
模型参数	C	[0.1, 1, 10]	控制正则化强度，调节欠拟合与过拟合之间的平衡
	gamma	['scale', 'auto']	用于非线性核（尽管这里主要使用 linear 核，仍保留 gamma 以做完整对比）
参数调优	网格搜索交叉验证折数	5	提高调参结果稳定性与泛化能力
模型评估	随机种子列表	[0, 42, 100, 200]	多随机状态验证结果稳健性，取平均准确率与标准差做整体性能评估

(2) 模型训练

在模型训练阶段，首先使用 TF-IDF 向量器提取文本特征，采用 1-2 元语法 (unigram 和 bigram) 以捕捉更丰富的上下文信息。TF-IDF 可通过公式 (4) 求得：

$$TF-IDF(t,d) = TF(t,d) \times \log\left(\frac{N}{DF(t)}\right) \quad (4)$$

其中， $TF(t,d)$  表示词项  $t$  在文档  $d$  中的词频， $DF(t)$  表示包含词项  $t$  的文档数量， $N$  是文档总数。

接着，采用支持向量机中的线性核函数 (Linear Kernel) 作为分类器核心，通过网格搜索 (GridSearchCV) 和 5 折交叉验证选择最优超参数。SVM 旨在构造一个最优的超平面，使得正负样本间的间隔最大，其线性核的决策函数表达式为公式 (5)：

$$f(x) = w^T x + b \quad (5)$$

其优化目标如下：

$$\min_{w,b} \frac{1}{2} \|w\|^2 \quad \text{subject to } y_i (W^T X_i + b) \geq 1 \quad (6)$$

其中， $w$  为权重向量， $b$  为偏置项， $y_i \in \{-1,+1\}$  为样本标签。

2.1.3 模型性能分析

TextCNN 和 SVM 模型在情感二分类任务中都表现出良好的性能，整体准确率达到 86% 和 83%，如表 6 所示。从两种模型的正向与负向标签的精确率、召回率和 F1 值看，模型在两类样本上的分类能力较为均衡。宏平均与加权平均的各项指标进一步证明模型在处理样本分布相对均衡的情况下具有较强的泛化能力。这一结果说明，TextCNN 在特征自动学习与语义建模方面具有一定优势；但同时也验证了在特征工程优化良好的前提下，SVM 依然具备较强的分类性能与泛化能力。整体上，两种模型的性能均具实用价值，TextCNN 略优于 SVM，适合用于情感表达复杂的评论文本场景。

表 6 TextCNN 和 SVM 模型性能比较

模型	项目	精确率	召回率	F1 值	样本数量
TextCNN	标签 0 (负向)	0.88	0.83	0.86	1171
	标签 1 (正向)	0.83	0.89	0.86	1121
	准确率			0.86	2292
	宏平均	0.86	0.86	0.86	2292
	加权平均	0.86	0.86	0.86	2292
SVM	标签 0 (负向)	0.86	0.80	0.83	1171
	标签 1 (正向)	0.80	0.86	0.83	1121
	准确率			0.83	2292
	宏平均	0.83	0.83	0.83	2292
	加权平均	0.83	0.83	0.83	2292

TextCNN 模型的 AUC 为 0.93，SVM 模型为 0.91，表明两者均具备较强的分类能力，但前者略优，如图 6 所示。ROC 曲线在不同阈值

下呈现较高的 TPR，且假阳性率控制良好，说明模型在处理正负样本分布不均的情况下仍具有较高的判别性能。

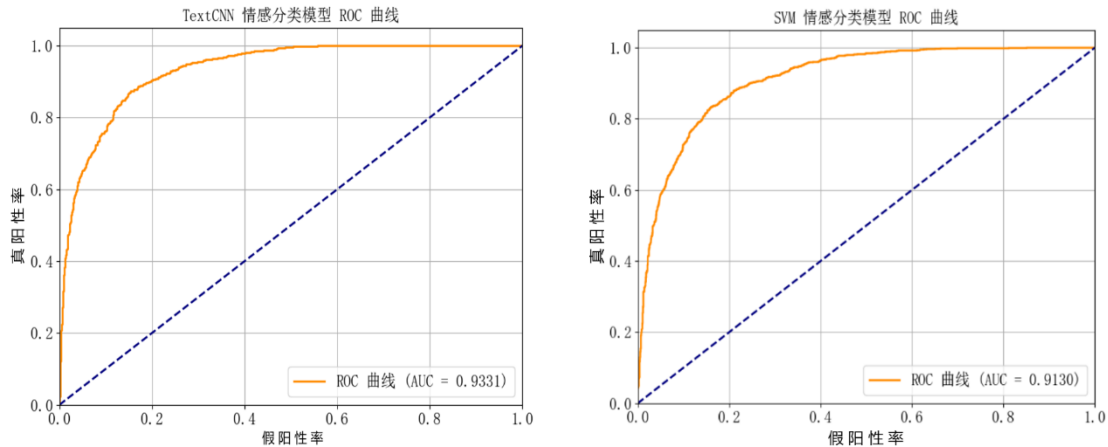


图6 TextCNN 模型的 ROC 曲线图(左)和 SVM 模型的 ROC 曲线图(右)

### 3 实验结果与分析

#### 3.1 实验结果评价指标

为评价 TextCNN 情感分类模型和 SVM 情感分类模型在中文情感分类任务中的表现,由研究人员人工核对其分类结果的正确性,统计其在情感分类结果的准确率、精确率、召回率及 F1 值,可通过公式(7)至(12)求得:

$$\text{准确率} = \frac{(\text{真阳性} + \text{真阴性})}{(\text{真阳性} + \text{真阴性} + \text{假阳性} + \text{假阴性})} \quad (7)$$

$$\text{精确率 (正向)} = \frac{\text{真阳性}}{(\text{真阳性} + \text{假阳性})} \quad (8)$$

$$\text{精确率 (负向)} = \frac{\text{真阴性}}{(\text{真阴性} + \text{假阴性})} \quad (9)$$

$$\text{召回率 (正向)} = \frac{\text{真阳性}}{(\text{真阳性} + \text{假阴性})} \quad (10)$$

$$\text{召回率 (负向)} = \frac{\text{真阴性}}{(\text{真阴性} + \text{假阳性})} \quad (11)$$

$$F1 \text{值} = \frac{2 * (\text{精确率} * \text{召回率})}{(\text{精确率} + \text{召回率})} \quad (12)$$

其中,真阳性表示被预测的数据为正向情感,实际也是正向情感;真阴性表示被预测数据的真实情感和预测情感都为负向;假阳性表示预测为正向实际为负向;假阴性表示预测为负向实际为正向。

#### 3.2 实验结果分析

将测试集的 1400 条评论分别输入训练好的 TextCNN 模型和 SVM 模型,获取和记录每条评论的情感分类结果。表 7 的情感分类结果显示,TextCNN 准确率达到 96%,SVM 为 93%,TextCNN 在此测试中表现更优。

表7 情感分类结果分析

模型	情感类型	数量	精确率	召回率	F1 值	准确率
TextCNN	正向	714 条	0.95	0.97	0.96	0.96
	负向	686 条	0.97	0.95	0.96	
SVM	正向	763 条	0.90	0.98	0.94	0.93
	负向	637 条	0.98	0.89	0.93	

### 3.3 错误样本分析

模型的错误分类主要分为两类，如表 8 所示。其中，第一类是实际情感为正向但被误判为负向的样本，如表 9 所示；第二类是实际情感为负向但被误判为正向的样本，如表 10 所示。总体而言，两种模型对负向情感的误判明显多于对正向情感的误判，表明模型在负向情绪识别方面面临更大挑战。从模型性能来看，TextCNN 的整体误判数量（56 条）低于 SVM（91 条），尤其是在负向情感被误判为正向的情况下，TextCNN（35 条）显著优于 SVM（77 条），体现了 TextCNN 在处理复杂语义结构和负面情绪表达方面的优势；而在实际正向情感被预测为负向的情形下，SVM 误判略低。尽管 TextCNN 在整体准确率上领先，但在识别部分表达平淡的积极情感时，SVM 表现出了更强的稳定性。

表 8 错误样本统计

模型	实际为正预测为负	实际为负预测为正	合计
TextCNN	21 条	35 条	56 条
SVM	14 条	77 条	91 条

针对第一类错误的分析发现，这类评论通常表达积极，但语言相对平淡或使用大量中性描述，如“电梯速度快，电梯安全，电梯空间大，无异味很放心，服务好”。尽管整体评价正向，但缺少显著的情绪标识词汇，使模型难以准确捕捉整体积极情感。此外，评论内容复杂多维度时，例如“酒店位置很好离春熙路很近，房间隔音效果很好，停车场就在小院很方便，电视机可以投屏，早餐味道很不错餐厅环境舒服，服务很好”，模型容易受语义复杂性干扰，对整体情绪判断失误。比较 TextCNN 与 SVM 的表现发现，TextCNN 在识别复杂语境和隐性积极情感方面略优于 SVM，体现出深度学习模型更强的语义建模能力。

表 9 实际为正向被误判为负向的样本

模型	序号	预测	实际	被预测文本
TextCNN	01	负向	正向	电梯速度快，电梯安全，电梯空间大，无异味很放心，服务好。
	02	负向	正向	酒店位置很好离春熙路很近，房间隔音效果很好，停车场就在小院很方便，电视机可以投屏，早餐味道很不错餐厅环境舒服，服务很好。
	03	负向	正向	位置方便很好找，离地铁站很近交通方便，酒店周围很多吃的很不错，前台小哥哥小姐姐很热情服务专业，有帮忙升级房型房间，干净卫生很安静没有噪音，床垫很舒服办理入住也很快。
SVM	01	负向	正向	电梯速度快，电梯安全，电梯空间大，无异味很放心，服务好。
	02	负向	正向	酒店位置很好离春熙路很近，房间隔音效果很好，停车场就在小院很方便，电视机可以投屏，早餐味道很不错餐厅环境舒服，服务很好。

针对第二类错误分析表明，负面评论被误判为正向通常涉及情绪表达转折、否定和讽刺语气，如“体验感不好，原房续订还要加钱你们就缺这几十块钱吗”，尽管语言具有明显的负向表达，但句中讽刺语气模型未能有效识别。

再如，“虽然说赠送早餐就写好评，但是因为不愉快入住不得不如实评价”，此类具有明显情绪转折结构的评论，模型亦易误判为正向。在对比分析 TextCNN 和 SVM 的误判情况时发现，两种模型对否定句式和情绪转折的敏感性

均有限，但 TextCNN 对讽刺与隐晦表达的识别略强于 SVM。这说明未来需进一步优化两种模

型，特别是在复杂语言现象的理解能力方面，以提升整体分类效果。

表 10 实际为负向被误判为正向的样本

模型	序号	预测	实际	被预测文本
TextCNN	01	正向	负向	体验感不好,原房续订还要加钱你们就缺这几十块钱吗,携程定的不是这个房型吗,价钱不要提前说吗,员工基本的常识都没有吗,影响心情。
	02	正向	负向	交通还可以,卫生好差,房间好小,环境好差,服务也不好,前台一直叫我们扫什么二维码。
	03	正向	负向	都退房了正休息呢来打电话要五星好评吓我一跳,我让你们升个级都要单独加钱才给升,最气人的是你加多点也行就十块钱,你加多点我都不那么气。房间沙发跟猫挠了一样才两年的酒店设施也不知道维护不会再去。
	04	正向	负向	虽然说赠送早餐就写好评,但是因为不愉快的人住不得不如实评价并希望给予和我们一样旅行中的朋友们正确的指引, ..., 当天下飞机就很疲惫被相同规律的声音折磨不堪……
	05	正向	负向	跟图片一样又不一样,比看图小太多了,不喜欢好小的小酒店啥都小小的门都是小小的,妈呀太难受了住得双床房也很拥挤反正不好。
SVM	01	正向	负向	当时不开发票,后补发票竟然还邮寄到付,这么差钱了吗?
	02	正向	负向	酒店给安排了间出电梯口的 1102 房间,刚开始没多想隔音咋样就没换房间,没想到到了晚上吵死了,客人进进出出都听的到。这还没啥,但酒店有那个自动机器人应该是送外卖一出电梯因为不平铛的一声直接整的一晚上没睡着气死了。
	...	...	...	...
	10	正向	负向	服务还行,房间很小很闷热死了,这是我住过的房间最小的酒店,昨天酒店说天气不稳定没开冷空调给我一个风扇结果我一夜没睡着以后不可能来这家了。
	11	正向	负向	跟图片一样又不一样,比看图小太多了,不喜欢好小的小酒店啥都小小的门都是小小的,妈呀太难受了住得双床房也很拥挤反正不好。

## 4 结论与展望

本文以中文酒店在线评论为研究对象,构建了 TextCNN 深度学习模型与 SVM 传统机器学习模型,进行了情感二分类的系统实证对比分析。实验结果表明,TextCNN 与 SVM 模型在测试数据中的分类准确率分别达到 96% 和 93%,表现出较好的泛化能力和实用价值。其中,TextCNN 模型在训练数据中准确率略高于 SVM,说明其在自动学习复杂语义特征方面具有一定优势,尤其在处理负面情绪表达的复杂语境中表现更为稳健。

本文也发现了模型的局限性:一方面,研究局限于正负二分类情感任务,未涉及更为精

细的情感强度分级或多类别情感分析,限制了模型的应用场景;另一方面,训练数据量相对有限,可能影响模型在实际大规模评论场景中的泛化能力。此外,模型对中性情感表达及含蓄语言的识别能力仍需进一步提升。

未来的研究将从以下两个方向深入:一是利用更先进的预训练语言模型(如 BERT、ERNIE 等)来增强对评论文本语义的理解,特别是针对中文多情感字符级信息的捕捉能力<sup>[15-17]</sup>;二是尝试将预训练模型与深度学习架构(如 ERNIE、LSTM)进行深度融合,以进一步提升在复杂酒店评论语境下的分类精度<sup>[18]</sup>。从而更好地满足酒店服务管理与用户满意度评估的实际需求。

## 参考文献

- [1] 李杨, 徐泽水, 王新鑫. 基于在线评论的情感分析方法及应用 [J]. 控制与决策, 2023, 38(2): 304-317.
- [2] GHARZOULI M, HAMAMA A K, KHATTABI Z. Topic-based sentiment analysis of hotel reviews[J]. Current Issues in Tourism, 2021, 25(9): 1368-1375.
- [3] BIRJALI M, KASRI M, BENI-HSSANE A. A comprehensive survey on sentiment analysis: Approaches, challenges and trends[J]. Knowledge-Based Systems, 2021(226): 107134.
- [4] Grljević O, Bošnjak Z, Kovačević A. Opinion mining in higher education: a corpus-based approach[J]. Enterprise Information Systems, 2022, 16(5): 1773542.
- [5] 张财, 马自强, 闫博. 基于机器学习的政务微博情感分析模型设计 [J]. 计算机工程, 2024, 50(12): 386-395.
- [6] 于海涛, 刘竞泽, 唐尧杰. 基于情感词典的民宿评论情感分析——以桂林阳朔为例 [J]. 重庆交通大学学报 (社会科学版), 2025, 25(2): 80-89.
- [7] 张小艳, 白瑜. 基于加权融合字词向量的中文在线评论情感分析 [J]. 计算机应用研究, 2022, 39(1): 31-36.
- [8] YOUNG T, HAZARIKA D, PORIA S, CAMBRIA E. Recent trends in deep learning based natural language processing[J]. IEEE Computational Intelligence Magazine, 2018, 13(3): 55-75.
- [9] KIM Y. Convolutional neural networks for sentence classification[J]. Transactions of the Association for Computational Linguistics, 2014(2): 175-187.
- [10] 张鑫玉, 才智杰. 基于 Bert-BiGRU-CNN 的文本情感分析 [J]. 计算机仿真, 2023, 40(7): 519-523.
- [11] 龙宇, 李秋生. 基于 Self-Attention 和 TextCNN-BiLSTM 的中文评论文本情感分析模型 [J]. 石河子大学学报 (自然科学版), 2025, 43(1): 111-121.
- [12] 刘逸, 肖文杰, 关僖, 等. 中文在线旅游评论的语言积极性偏见研究 [J]. 旅游科学, 2024, 38(11): 48-66.
- [13] 史达, 王乐乐, 衣博文. 在线评论有用性的深度数据挖掘——基于 TripAdvisor 的酒店评论数据 [J]. 南开管理评论, 2020, 23(5): 64-75.
- [14] 惠调艳, 王智, 何振华, 等. 基于词典-TextCNN-Word2Vec 组合模型的在线评价细粒度情感分析 [J]. 情报理论与实践, 2025, 48(2): 168-177.
- [15] MALIK S S, ILYAS M, HAQ Y U, et al. Multi-modal emotion detection and sentiment analysis[J]. IEEE Access, 2025(13): 59790-59810.
- [16] 王董祺, 杨洪山, 黄勃, 等. 基于 BERT 的面向多情感字符级信息的文本情感分类方法 [J]. 武汉大学学报 (理学版), 2025, 71(4): 506-516.
- [17] 蒲秋梅, 黄方俐, 王辉. 基于 BERT-LSTM 模型的情感分析研究 [J]. 中国电子科学研究院学报, 2023, 18(10): 912-920.
- [18] 齐梦娜, 朱丽平, 李宁. 基于 ERNIE 和 CNN 的在线评论情感分析模型 [J]. 计算机应用, 2022, 42(S1): 7-11.

(责任编辑: 许德山)