

基于SenticNet的电商客服对话细粒度情感波动分析

朱陈云婧, 马彪*

东华大学旭日工商管理学院, 上海

收稿日期: 2024年12月8日; 录用日期: 2025年1月1日; 发布日期: 2025年1月8日

摘要

电商的迅速发展为消费者带来了便捷,但客户服务面临的挑战也日益增多。客服响应效率与问题解决效果对用户体验至关重要,然而目前广泛采用的事后评价机制无法对用户情绪进行即时监控,难以及时识别和处理用户情绪波动,限制了服务满意度的提升。为克服这一不足,文章提出了一种基于SenticNet情感词辅助的跨句上下文对话情绪分类模型(SLED-EC)。该模型不仅整合了词语级与句子级的情感识别任务,以更精细地捕捉情感细节,还通过分析上下文信息来识别情绪的转变点,构建情感波动曲线,动态捕获多轮对话中的情感变化。实验结果表明,SLED-EC在准确率(67.78%)、召回率(83.81%)和F1值(72.36%)方面优于基准模型,展示了其在客服对话场景中情感分析的优势。该研究为电商客服系统提供了一种新的解决方案,有助于提升用户满意度和优化服务体验。

关键词

对话情感分析, 上下文依赖, 情绪波动曲线, 电商客服

Fine-Grained Emotion Fluctuation Analysis of E-Commerce Customer Service Dialogues Based on SenticNet

Chenyunjing Zhu, Biao Ma*

Glorious Sun School of Business and Management, Donghua University, Shanghai

Received: Dec. 8th, 2024; accepted: Jan. 1st, 2025; published: Jan. 8th, 2025

*通讯作者。

文章引用: 朱陈云婧, 马彪. 基于 SenticNet 的电商客服对话细粒度情感波动分析[J]. 服务科学和管理, 2025, 14(1): 24-36.
DOI: 10.12677/ssem.2025.141004

Abstract

The rapid advancement of e-commerce has significantly enhanced consumer convenience, yet it has simultaneously exacerbated the challenges faced by customer service. Customer service response efficiency and problem-solving effectiveness are critical to the user experience. However, widely adopted post-feedback mechanisms are inadequate for real-time monitoring of user emotions, hindering timely identification and management of emotional fluctuations, which limits improvements in service satisfaction. To address this shortcoming, this paper proposes a SenticNet lexicon-assisted cross-sentence contextual dialogue emotion classification model (SLED-EC). The proposed model integrates both word-level and sentence-level emotion recognition tasks, enabling more precise capture of emotional nuances. Additionally, it analyzes contextual information to identify emotional transition points and constructs emotional fluctuation curves, dynamically capturing emotional changes in multi-turn dialogues. Experimental results demonstrate that SLED-EC significantly outperforms baseline models in accuracy (67.78%), recall (83.81%), and F1-score (72.36%), underscoring its advantages in emotion analysis for customer service dialogue scenarios. This study offers a novel solution for e-commerce customer service systems, contributing to improved user satisfaction and optimized service experience.

Keywords

Dialogue Sentiment Analysis, Cross-Sentence Context, Emotion Fluctuation Curve, E-Commerce Customer Service

Copyright © 2025 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

为应对不断增长的用户需求, 淘宝、京东等电商平台采用在线客服系统提供服务。用户满意度是衡量服务质量的重要指标, 因此电商平台建立了服务评级系统, 通过用户评价对在线客服的服务质量进行评级。然而, 目前普遍采用的评级系统多为事后评价, 即服务结束后用户才能进行反馈, 导致客服无法在对话过程中及时调整服务, 进而引发用户负面情绪升级和流失。

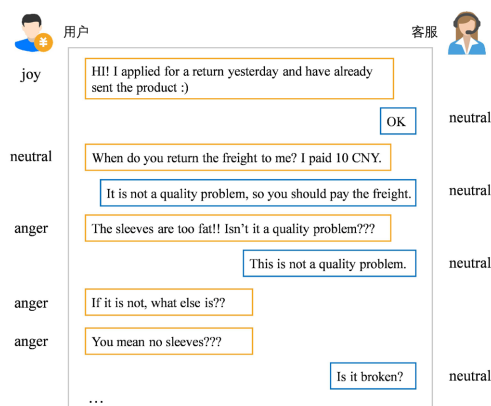


Figure 1. A service dialogue in E-commerce

图 1. 一个电商客服的服务对话示例

在线客服对话是一种连续交互过程, 包含用户丰富的情绪变化。如图 1 所示的对话中, 用户一开始情绪积极, 并使用“:)”表达友好, 但随着对话的进行, 客服反复强调 “It is not a quality problem” 并要求用户承担退货运费, 忽视了用户的反馈, 导致用户情绪转为愤怒和不满。用户情绪能够反映对客服服务的满意度, 而消极的情绪可能影响复购率和品牌忠诚度。因此, 理解并分析用户情绪的变化, 对于提升客服服务质量和用户体验至关重要。

情感分析(Sentiment Analysis)作为自然语言处理(NLP)中的核心任务之一, 旨在识别文本中的情感倾向。与词语级[1]、方面级[2]的情感分析相比, 对话情感分析因其多轮互动和上下文依赖性更具挑战性。对话中的情感往往跨越句子, 并随着对话进程不断变化, 这增加了情绪追踪的难度。单句分析往往无法准确捕捉情感的动态变化, 因此结合上下文信息[3][4]进行多轮对话的情感建模至关重要。然而, 现有的对话情感识别方法多集中于为单句情感标签分配[5], 未能充分考虑对话情感的连续性和全局性。在多轮对话中, 情绪波动曲线能够更直观地展示情绪的全局趋势[6], 弥补逐句分析的不足。对话中的情感常由两个或多个参与者共同影响, 因此理解不同发言者间的情感互动和相互反应同样重要[7]。与极性分析相比, 对话情感分析应关注更细粒度的情绪, 如惊讶、失望和愤怒等。此外, 特定对话场景中常有高频词汇, 如客服对话中的“发货”、“退款”、“质量问题”等, 深入理解这些词语的语境含义有助于更准确地识别对话情感。

为应对这些挑战, 本文提出了一种基于情感词辅助的跨句上下文对话情绪分类模型(Sentiment-Lexicon Enhanced Cross-Sentence Contextual Dialogue Emotion Classification Model, SLED-EC)。首先结合 SenticNet 情感词典和 RoBERTa 模型, 构建了包含词语级和句子级情感标签的训练数据集。接着, 通过 [CLS] token 将对话中的各语句连接为长序列, 并使用 Character Embeddings 替代 BERT 的 Segment Embeddings, 使模型更好地区分不同发言者的情绪和语义差异, 增强对多轮对话情感的理解。最后, 通过分析对话中单句的情感分布, 构建整个对话的情绪波动曲线。SLED-EC 的主要创新点如下: 1) 结合 SenticNet 实现词语级情感识别, 增强模型对细粒度情感变化的捕捉能力; 2) 利用跨句上下文信息, 准确识别多轮对话中的情感转折与动态变化; 3) 构建情绪波动曲线, 提供用户情感状态的全局视角。该模型在电商平台客服系统中具有广泛的应用潜力, 可以通过监测和分析用户的情绪波动, 辅助客服及时调整服务, 预防负面情绪的积累, 提升用户满意度。此外, 模型还可应用于智能对话系统, 通过精确的情绪识别提供个性化和贴心的服务体验。

2. 相关研究

2.1. 上下文依赖的对话情感分析

对话情感通常具有上下文依赖性, 每个句子的情感不是孤立存在的, 这种依赖性增加了多轮对话情感分析的复杂性。早期的情感分析研究主要集中于单句或短文本[1]的情感分类, 未能充分考虑上下文因素的影响。然而, 在多轮对话中, 单个句子的情感往往不足以反映对话的整体情感走向, 因为情感状态会随着对话进展而动态变化。

长短期记忆网络(LSTM)和 Transformer 等模型因其对序列数据的处理能力, 被广泛应用于捕捉对话中的上下文信息。LSTM 能够有效地保留对话中的长距离依赖[8], 而 Transformer 的自注意力机制能够在更大范围内捕捉情感变化的细节[9]。一些研究将 LSTM 模块与 Transformer 结合, 以进一步提升对对话中上下文信息的建模能力[10]。随着技术的发展, 多轮对话情感分析逐渐从单句分析向上下文依赖的动态分析转变。

随着对话系统的复杂性增加, 模型不仅需要捕捉显性情感信息, 还需理解情感的动态转变, 这对模型的上下文处理能力提出了更高的要求。在客服满意度[11]和对话响应[12]等应用中, 基于上下文的多轮

对话情感分析被证明在提升用户体验和理解用户意图方面发挥重要作用。此外, 对话情感分析还应关注情感的连续性和整体趋势, 其中情绪波动曲线作为一种整体建模方法, 能够动态反映用户情绪在整个服务过程中的变化, 帮助客服系统更全面地识别用户情绪变化的关键时刻。在电商客服场景中, 情绪波动曲线能为客服提供情绪变化的趋势参考, 尤其当用户情绪波动剧烈时, 系统可以为客服提供适当的沟通策略建议或建议更换客服, 以更有效地管理用户情绪, 提升服务质量。

2.2. 基于 BERT 的对话情感分析

Devlin 等[13]于 2018 年提出了基于 Transformer 编码器的 BERT 模型。该模型采用了 MLM (Masked Language Model)和 NSP (Next Sentence Prediction)预训练任务, 能够学习融合左右上下文信息的深层双向语言表示。在 BERT 模型中, 每个句子在开头插入 CLS 标记, 末尾插入 SEP 标记, 将文本转换为字向量、文本向量和位置编码之和作为输入, 如图 2 所示。

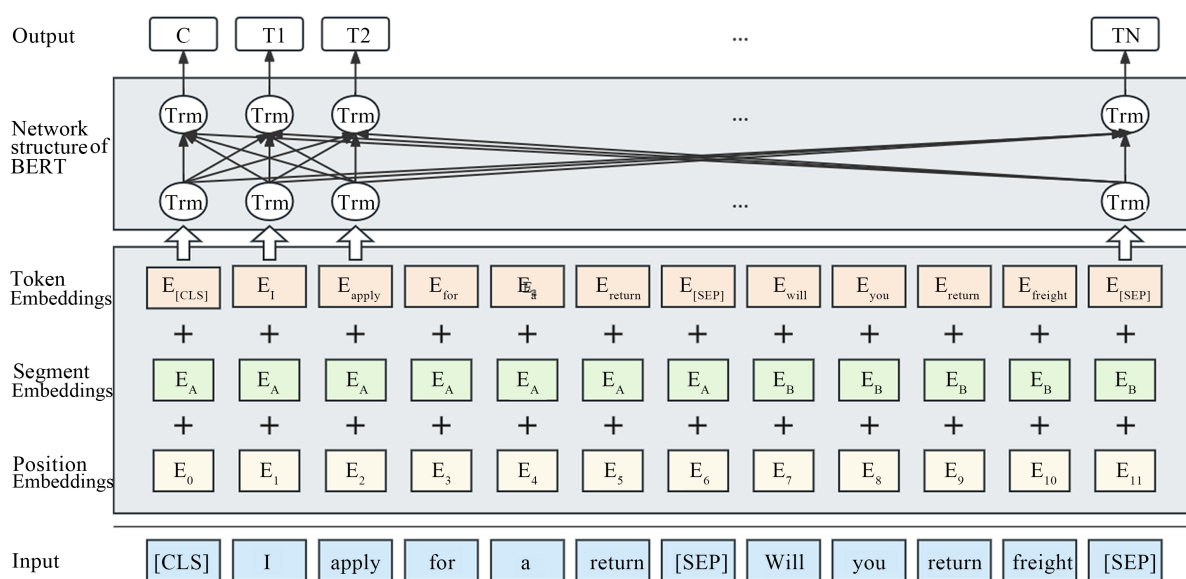


Figure 2. Architecture of BERT

图 2. BERT 模型结构

近年来, BERT 模型在自然语言处理任务中取得了显著进展, 尤其在情感分类任务中表现优异[14]。与传统的单向语言模型相比, BERT 不仅能够识别文本的整体情感倾向, 还能捕捉更为细腻的情感变化, 特别是在处理情感转折点和细粒度情感类型(如愤怒、爱、喜悦等)方面表现卓越[15]。

对话情感分析涉及多轮互动和复杂上下文关系, BERT 因其强大语境理解能力而广泛应用于对话系统。Zhang 等[16]提出 Dialogue BERT, 通过改进 BERT 的输入表示并引入对话特征, 增强了模型在对话理解和情感分析任务中的表现。Yang 等[17]基于 BERT 提出了一种情感动态建模方法, 用于在多轮对话中捕捉并预测情感的时序变化, 从而提高了对话系统的情感分析精度和上下文理解能力。BERT 不仅能够捕捉对话中的细粒度情感, 还能为客户服务、对话生成等下游任务提供了强有力的支持[18]。

3. 客服对话情绪识别模型

3.1. 问题描述

为挖掘用户在电商客服服务过程中的动态情绪体验, 本文将以对话中句子级的情绪为研究对象。一

句话可能包含多种不同的情绪, 如当用户希望商家对退货运费做出售后时(如图 1): “If it is not, what else is?? You mean no sleeves???” , 句子表达了用户的惊讶和愤怒, 且愤怒相对惊讶更高。因此模型不仅需要识别句子中是否有情绪信息, 还需要识别具体情绪的强度大小。此外, 客服对话通常是一个多轮交互过程, 包含了多个语句, 又可以进一步实现情绪的动态波动建模。

如图 3 所示, SLED-EC 模型将基于词语级、句子级两个情绪分类任务展开, 通过对语句中词语情绪的预测提升模型对语句情绪的理解。一段对话由 N 个句子组成, 即对话 $dialogue = [s_0^{spk_0}, s_1^{spk_1}, \dots, s_{N-1}^{spk_{N-1}}]$, 其中上标 spk_t 表示 t 时刻语句对应的说话人, $spk_t \in \{\text{顾客, 客服}\}$, $s_t^{spk_t}$ 表示 t 时刻 spk_t 的说话内容。每个语句经过分词处理得到词语 $word \in w_t$, 即 t 时刻的句子 $s_t^{spk_t} = [w_0, w_1, \dots, w_{M-1}]$ 。词语 w_t 和句子 s_t 将分别具有对应的情绪标签 $Label^w$ 和 $Label^s$ 。由于对话中通常存在上下文语境关系, 因此除了考虑 t 时刻的句子 s_t , 还需要考虑 t 时刻前的语句 $[s_0, s_1, \dots, s_{t-1}]$, 甚至 t 时刻后的语句 $[s_{t+1}, s_{t+2}, \dots, s_{N-1}]$ 。

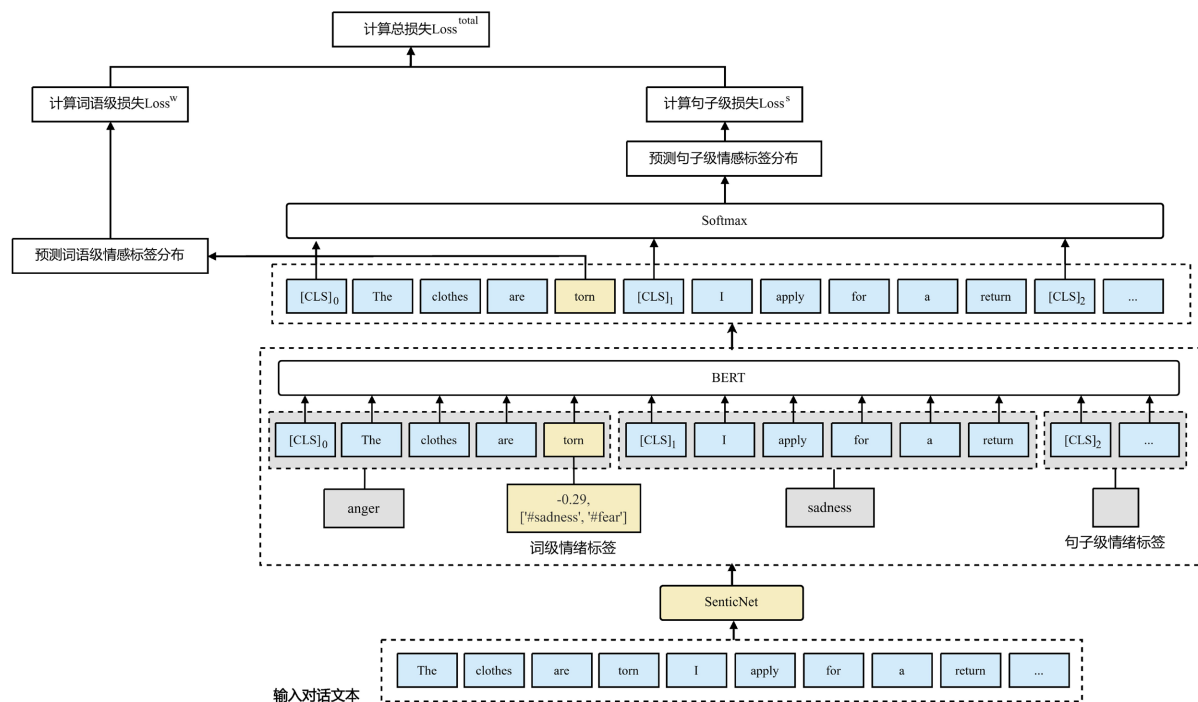


Figure 3. Structure of SLED-EC model

图 3. SLED-EC 模型结构

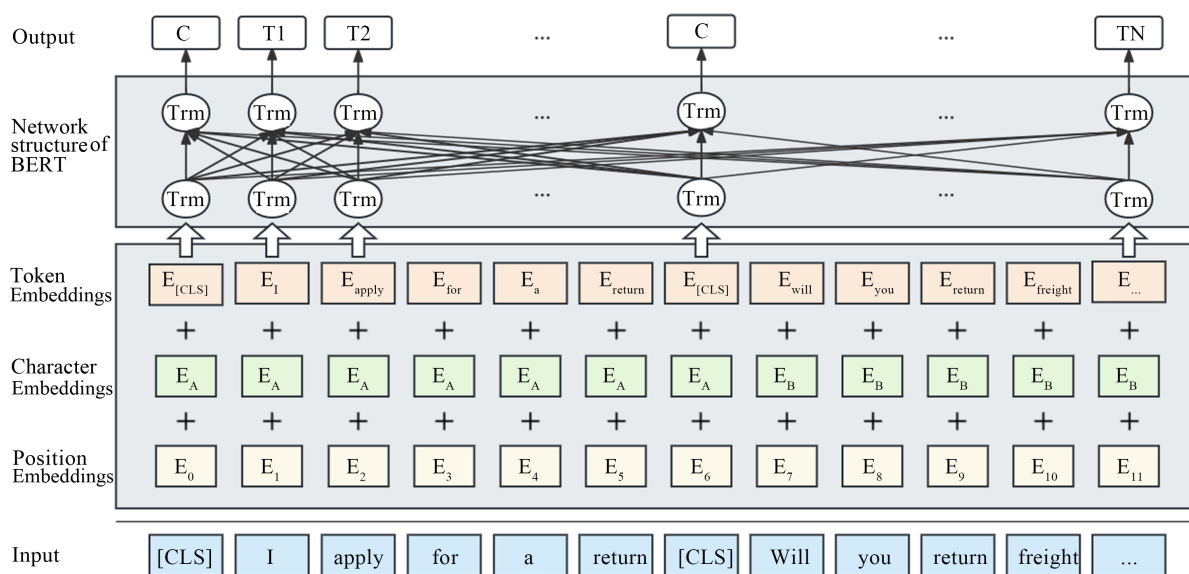
3.2. 模型构建

情感分析任务中, 外部情感知识可以看作是文本语义的情感补充来源。SenticNet [19] 是一个公开的情感知识库, 结合了情感和认知两个维度来理解和分析文本数据, 并为每个概念提供相应的情感分数, 被广泛应用到情感分析、情感标签生成等任务中。本文利用 SenticNet 情感词典构建词级训练数据集: 对话中句子的词语 w_t 的情绪标签 $Label^w \in \{\text{joy, pleasantness, eagerness, calmness, anger, sadness, disgust, fear}\}$, 及对应的情感分数 $SenticNet(w_t) \in [-1, 1]$ 。强消极词的情感分数接近 -1 , 强积极词情感分数接近 1 , 当 $SenticNet(w_t) = 0$ 时, 表示词语 w_t 的情感倾向为中性或者在词典中不存在。SenticNet 中词语的情绪标签例如表 1 所示。本文直接采用 MELD 对话数据集[20]中的情绪分类: anger、disgust、sadness、joy、neutral、surprise 和 fear, 即句子 s_t 具有情绪标签 $Label^s \in \{\text{anger, disgust, sadness, joy, neutral, surprise, fear}\}$ 。

Table 1. Examples of word sentiment score and mood tags in SenticNet**表 1.** SenticNet 中词语情感分数及标签示例

单词 w_i	$SenticNet(t_i)$	Mood tags
excellent	0.744	#joy, #eagerness
fantastic	0.87	#joy, #calmness
damage	-0.45	#sadness, #fear
awful	-0.94	#sadness, #disgust

对话中语句的情绪强度分布可以视为一组基于 BERT 的上下文语境建模问题。将带有情绪标签 $Label^{*}spk$ 的语句通过 [CLS] token 连接后, 形成 “[CLS] I apply for a return [CLS] Will you return freight [CLS]...” 的长序列输入至 BERT 模型中, 并转换为 Token Embeddings、Character Embeddings、Position Embeddings 之和作为 BERT 模型的最终输入(如图 4 所示)。需要注意的是, Character Embeddings (即 Customer Embeddings 与 Service Embeddings)取代了原 BERT 模型中的 Segment Embeddings, 用于识别各语句的说话人。模型将预测得到对话中每个语句的情绪标签及分布, 通过权重计算得到相应的情感分数, 最终生成不同对话的语句间情感波动情况。

**Figure 4.** Architecture of BERT in SLED-EC model**图 4.** SLED-EC 中 BERT 的模型结构

BERT 模型由多个 Transformer 层重叠而成, Transformer 层主要由 Encoder 和 Decoder 组成。其中, Encoder 层包含一个自注意力网络层和一个前馈神经网络, 而 Decoder 层比 Encoder 层多一层注意力层。在 Transformer 架构中的注意力机制中, Query 用于衡量与每个 Key 的相关性, Key 用于构建注意力得分的计算基础, 而 Value 用于根据注意力得分进行加权求和。多头注意力将信息进行融合生成最终的注意力输出:

$$MultiHead(Q, K, V) = Concat(Head_1, Head_2, \dots, Head_h)W^O \quad (1)$$

$$Head_i = Attention(QW_i^Q, KW_i^K, VW_i^V)$$

模型选择交叉熵损失(Cross-Entropy Loss)作为模型的主要损失函数:

$$Loss^w = -\sum_{i=0}^{S^w} \sum_{j=0}^{C^w} y_i^{j^w} \log(y_{pi}^{j^w}) \quad (2)$$

$$Loss^s = -\frac{1}{N} \sum_{i=0}^{S^s} \sum_{j=0}^{C^s} y_i^{j^s} \log(y_{pi}^{j^s}) \quad (3)$$

$$Loss^{Total} = Loss^w + Loss^s \quad (4)$$

其中, $Loss^{Total}$ 为任务最终的损失函数, $Loss^w$ 、 $Loss^s$ 分别表示词语级与句子级情感分类任务的损失函数。 S^w 、 S^s 是训练样本, C^w 、 C^s 分别表示两类任务的情绪类别数。 $y_i^{j^w}$ 、 $y_i^{j^s}$ 为训练数据中的真实标签, $y_{pi}^{j^w}$ 、 $y_{pi}^{j^s}$ 为预测标签, N 为对话中的语句数, 取均值以减少对话中语句数量差异造成的影响。最终得到语句的七类情绪概率分布。为进一步计算对话中每句语句的情感值, 本文对七类情绪标签进行了排序: joy、surprise、neutral、sadness、fear、disgust、anger, 并根据情绪的正负向和强度分配了相应的权重。通过加权计算得到语句的情感值为:

$$E_t = \sum_{j=0}^{C^s} w_j \times p_j(s_t) \quad (5)$$

其中, $p_j(s_t)$ 表示语句 s_t 中第 j 种情绪的概率分布, w_i 表示不同情绪的权重, 设定 $w_{joy} = 3$, $w_{surprise} = 1.5$, $w_{neutral} = 0$, $w_{sadness} = -1$, $w_{fear} = -1.5$, $w_{disgust} = -2$, $w_{anger} = -3$ 。

4. 实验结果及分析

4.1. 数据集

本文基于一个公开可用的电商领域服务对话数据集进行实验。该数据集[21]是京东公司在客服和客户的真实聊天数据上做了脱敏处理后的数据, 包含了共计 10,000 个对话数据。数据集中包含了每段对话的角色信息(0-用户, 1-客服)和具体对话内容, 由于对话文本属于短文本非结构化数据, 实验前将需要对文本进行预处理。文本预处理是文本挖掘的准备工作, 主要包括文本翻译、分词、去停用词等。EmoBERTa [22]是一种改进的 RoBERTa 模型, 支持多分类情感识别, 在两个对话情感分析数据集上表现优异。因此, 实验结合 EmoBERTa 和人工标注方法对在线客服对话数据集进行情感标注, 统计信息与标注示例如表 2、表 3 所示。数据集设定训练、开发与测试集的比例为 8:1:1。

Table 2. Statistics of emotion categories in the dataset

表 2. 数据集情绪类别统计信息

总句子数	joy	surprise	neutral	sadness	fear	disgust	anger
212,808	40,788	1,182	131,085	19,507	739	4,729	14,778

Table 3. Examples of sentences for emotion categories in the dataset

表 3. 数据集情绪类别句子示例

情绪类别	句子示例	情绪类别	句子示例
joy	Okay, thank you!	surprise	Really?
neutral	Can I pick it up myself?	sadness	I hope this issue can be resolved for me as soon as possible.
fear	Afraid I might not be back when the delivery arrives.	disgust	Someone I know bought this before and had the same problem.
anger	Are you sure I'll receive it by the [number x]???		

4.2. 实验设置及评价标准

本实验基于 BERT 模型进行对话情绪分类，由 12 个 Transformer 块、12 个自注意力头和 768 个隐藏单元组成，并使用预训练参数对 BERT 进行初始化，其他参数则随机初始化。采用 AdamW 优化器，设置初始学习率为 0.001， $\beta_1 = 0.9$ ， $\beta_2 = 0.999$ ，L2 正则化权重衰减为 0.01，学习率预热步数为 0，使用线性衰减学习率。所有结果为基于 5 次实验运行的平均值。

作为典型的分类问题，情绪分类的评价标准可以比照普通的分类问题处理，本文采用的评价指标是准确率(Precision)、召回率(Recall)和 F1 值，如下所示：

$$P = \frac{TP}{TP + FP} \quad (6)$$

$$R = \frac{TP}{TP + FN} \quad (7)$$

$$F1 = \frac{2 \times P \times R}{P + R} \quad (8)$$

P 和 N 表示模型实际的判别情况，T 和 F 评价模型的判别情况正确与否，TP、FP、TN、FN 分别表示真正例、假正例、真反例、假反例的样本数量。

4.3. 实验结果

4.3.1. 对比实验

为评估 SLED-EC 模型在客服对话情感识别任务中的表现，本文选择了以下几种对话情感分类模型作为基准，进行对比实验：

- 1) DMN [23]：动态记忆网络，通过动态存储和更新信息增强推理能力，尤其在长距离依赖的处理上表现优异。
- 2) BERT + MTL [24]：将情感识别与辅助说话者识别任务结合，提升对话情感识别的效果。
- 3) DialogueGCN [25]：利用图卷积网络处理对话中参与者之间的依赖关系，显著提升对话情感分析的效果。
- 4) DialogueBERT [16]：基于自监督学习的对话预训练编码器，通过改进 BERT 输入表示和引入对话特征，优化了对话情绪分类任务。

Table 4. Comparison of experimental results across methods

表 4. 对比实验结果

模型	准确率 P (%)	召回率 R (%)	F1 值(%)
DMN [23]	57.06	73.39	62.03
BERT + MTL [24]	61.87	78.58	67.05
DialogueGCN [25]	59.60	76.74	64.44
DialogueBERT [16]	65.12	81.85	70.18
SLED-EC*	67.78	83.81	72.36

表 4 比较了 SLED-EC 模型与其他四种基准模型在电商客服对话情感识别任务中的表现。SLED-EC 在准确率(67.78%)、召回率(83.81%)和 F1 值(72.36%)这三个关键指标上均优于对比模型。尤其是在召回率和 F1 值方面，SLED-EC 展现了在情感识别上的全面性和平衡性，表明了 SLED-EC 在细粒度情感捕捉

和上下文依赖性分析方面的优势。

与其他模型相比, SLED-EC 更适合处理客服对话中的情感识别任务。通过情感词典辅助的多任务结构, SLED-EC 能够更好地挖掘特定场景中词语的不同语义和情绪强度。结合词级别和句级别任务, SLED-EC 能够捕捉到更多细粒度情感变化, 并结合上下文信息准确理解情绪转折, 从而在整体性能上表现突出。尽管 DMN、BERT + MTL 和 DialogueGCN 在某些方面表现出色, 但在处理多轮对话和电商场景下的情绪交互时, 它们的局限性开始显现。SLED-EC 则通过有效整合上下文信息和情感词, 展现了更强的适应性。

在图 1 所示的对话示例中, 用户的情绪从最初的积极(joy)逐渐转为愤怒(anger), 表现出对客服回应的强烈不满。SLED-EC 模型能够准确捕捉这一情绪变化过程。例如, 当用户询问 “Isn’t it a quality problem???” 时识别出用户不满情绪的显著上升, 并生成情绪波动曲线, 直观展示用户情绪的整体变化趋势。与其他模型相比, SLED-EC 结合了细粒度的词级情绪信息和跨句的上下文理解, 提升了模型性能, 并通过识别关键情绪转折点, 为客服系统提供了有效的情绪管理支持。

4.3.2. 消融实验

消融实验评估了不同模型配置对情感分类性能的影响, 具体设置如下:

- 1) 单句子输入模型: 即不带对话上下文的单句子输入。将对话中的单个句子直接输入 BERT 模型进行情绪识别, 忽略上下文信息和词级别的情绪标签, 为后续实验提供基准。
- 2) 词级别与句子级多任务情感分类方法: 结合 SenticNet 情感词典, 在 BERT 模型中加入词语级情绪识别任务, 以验证词级任务是否能提升句子级情绪分类的性能。
- 3) 跨句子上下文信息的情感分类方法: 将对话中前后关联的句子作为长序列输入模型, 并通过 Character Embeddings 区别说话人, 验证上下文信息对情感理解的增强效果。
- 4) 情感词辅助的跨句上下文对话情绪分类模型(SLED-EC): 同时结合词语级情绪识别与上下文信息, 最大化利用对话中的语义与情感细节, 以提升模型的情绪分类能力。

Table 5. Ablation study results

表 5. 消融实验结果

句子级	词语级	上下文信息	准确率 P (%)	召回率 R (%)	F1 值(%)
√	×	×	60.56	74.24	65.20
√	√	×	62.74	77.66	67.26
√	×	√	64.36	80.01	69.11
√	√	√	67.78	83.81	72.36

实验结果如表 5 所示, 展示了模型在不同配置下的性能表现。当仅依赖句子级别的情感识别时, 模型的准确率、召回率和 F1 值分别为 60.56%、74.24%和 65.20%, 为所有配置中表现最差的情况。引入词语级别情感识别后, 准确率和 F1 值分别提升至 62.74%和 67.26%, 但提升幅度有限。这是因为词语级别情感识别虽然增强了对细微情感的捕捉能力, 但在多轮对话中无法充分捕捉上下文中的情感关联。

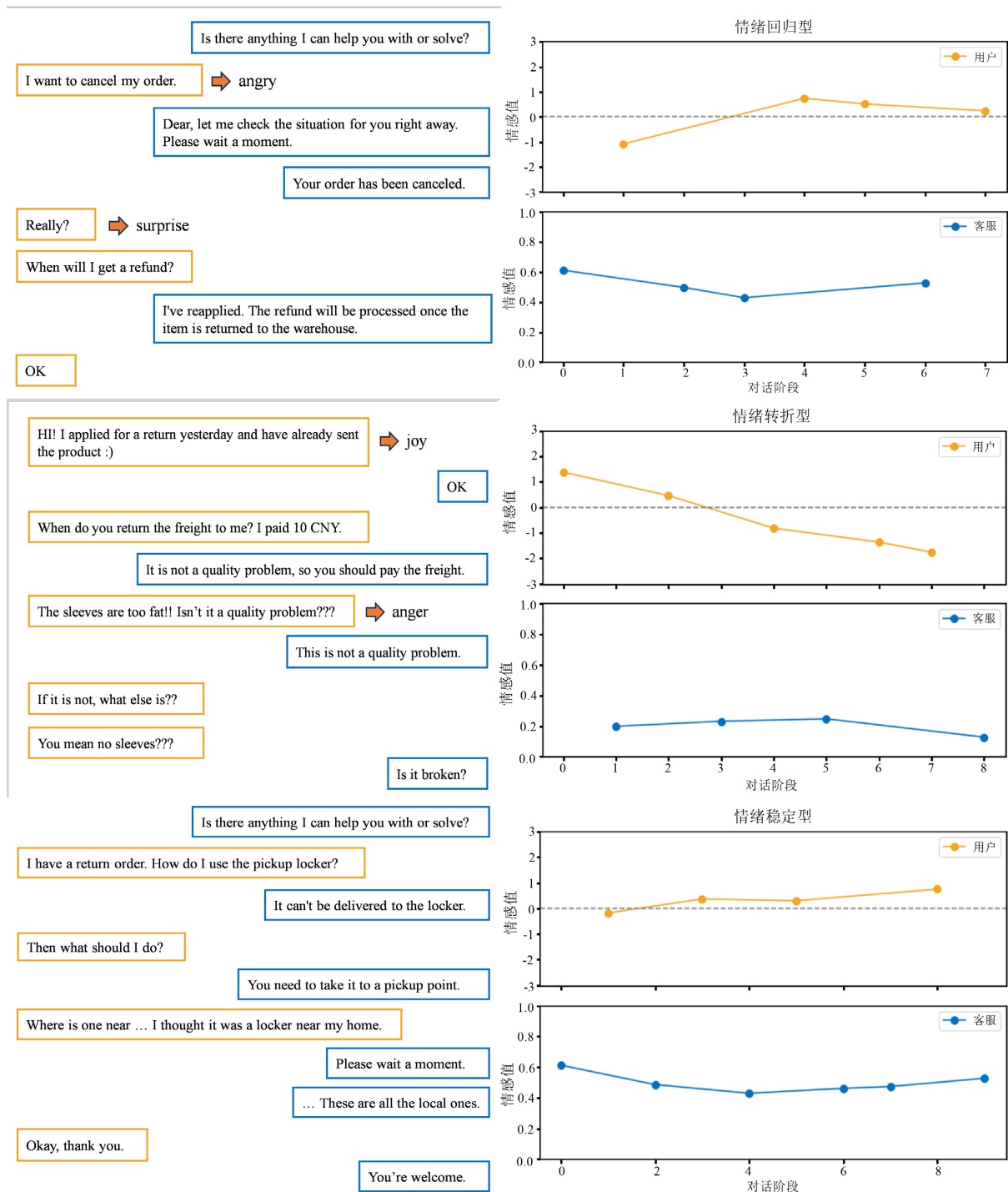
当进一步加入上下文信息时, 模型的准确率、召回率和 F1 值分别提升至 64.36%、80.01%和 69.11%。上下文信息的引入使模型能够理解对话中的前后关系, 更好地捕捉情感转折, 提升了情感分析的准确性。在多轮对话中, 用户情绪常因客服的回应逐步变化, 这种变化仅靠单句或词语级别分析难以捕捉, 需借助上下文信息来识别。

结合句子级、词语级情感识别及上下文信息的 SLED-EC 模型表现最佳, 其准确率、召回率和 F1 值

分别达到了 67.78%、83.81%和 72.36%，相比基础配置提升了 7.22、9.57 和 7.16 个百分点。多层信息的融合使 SLED-EC 能更精确地捕捉情感细节与情绪动态变化，不仅提高了整体性能，还增强了其在复杂情感交互场景中的适应性和灵活性。

4.3.3. 情绪波动曲线分析

如图 5 所示，四组客服对话的情绪波动曲线揭示了用户和客服的情感变化趋势。



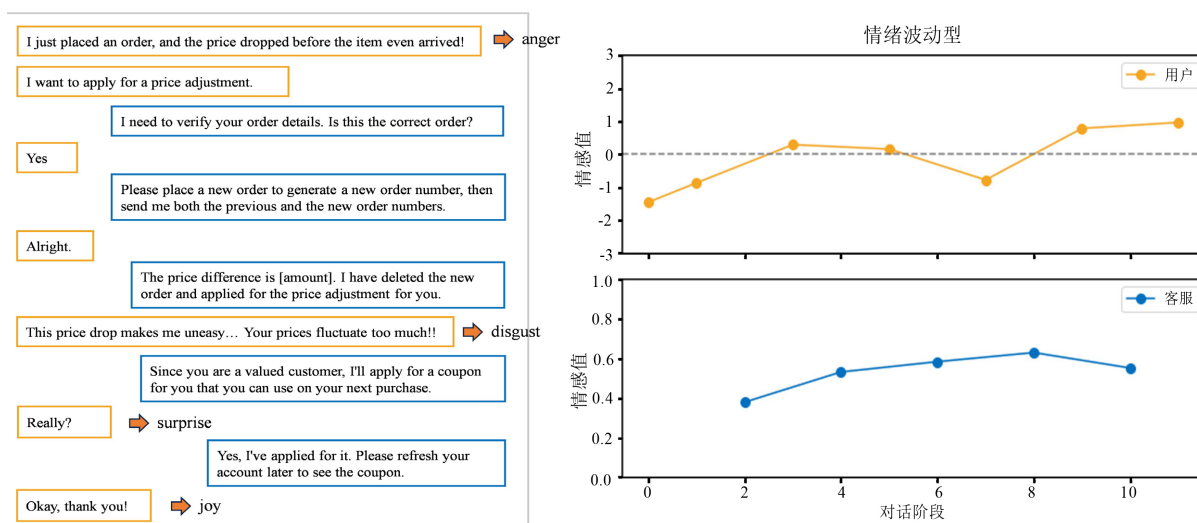


Figure 5. Dialogue content and emotional fluctuation curve

图 5. 对话内容与情绪波动曲线图

在对话过程中, 客服的情绪表达多为积极或中性, 而用户的情绪则更为多变。因此, 本文优先考虑用户的情绪, 并定义了以下四种情绪波动类型:

- 1) 情绪回归型: 用户起初表达出失望或愤怒等负面情绪, 在客服安抚下情绪逐渐回归至中立或积极状态。
- 2) 情绪转折型: 由于问题未能及时解决或服务未达用户预期, 用户情绪从积极转向负面。
- 3) 情绪稳定型: 用户情绪在对话过程中波动较小, 保持相对稳定。
- 4) 情绪波动型: 用户情绪在对话中多次变化, 出现显著的波动。

结合图 5 所示的对话内容和情绪波动曲线, 可以观察到在“情绪转折型”对话中, 因客服未能及时安抚或提供满意服务, 用户情绪从愉快(joy)转为愤怒(anger)。而在“情绪波动型”对话中, 用户情绪先得到了缓和, 但因平台或商家的降价促销活动导致未能满足预期, 再度出现反感情绪(disgust)。这种波动并非直接由客服的服务质量引起, 而是由于外部因素导致。通过监控用户情绪变化, 平台可以识别情绪转变的关键时刻, 从而优化客服沟通策略, 有效缓解用户不满并提升整体满意度。

5. 结论与展望

现有的对话情感分析方法在处理多轮对话的上下文语境以及电商服务场景下的细粒度情感分析方面, 仍存在一定的局限性。特别是在多轮对话中, 准确捕捉上下文关联和角色间的情感交互仍然是一个挑战。针对这一问题, 本文提出了一种结合 BERT 模型和 SenticNet 词典的情绪分类模型。通过整合词级和句子级的情感识别, 模型能够有效提升对话中每个语句的情感理解能力, 尤其在情感波动和角色情感交互复杂的场景中展现出优异表现。实验结果表明, 本文模型在客服对话情感分析任务中表现出较高的准确性, 尤其在情感变化检测和长对话情感预测方面具有明显优势。与现有方法相比, 本文模型不仅提升了对各类情感的识别精度, 还在情感波动场景下提供了更细致、更完整的分析, 为电商服务中的用户满意度评估提供了更有力的支持。

尽管 SLED-EC 模型在实验中取得了较好的效果, 但由于所使用的数据集主要来自于特定电商平台的客服对话, 这可能限制模型在其他平台或行业中的泛化能力。数据集中情感类别的分布不均, 可能导致模型在少数类情感(如“fear”或“disgust”)上的表现较弱。未来将扩展数据来源, 并结合多模态数据

(如语音、视频等), 以增强模型对对话情感的理解能力。考虑到智能客服系统对情感分析的实时性需求, 接下来将着重提高模型在动态对话中的响应速度和准确性。同时, 情感波动模式和趋势的预测也将成为研究的重点, 通过进一步优化模型, 推动智能客服系统向更加个性化的方向发展, 从而提升用户体验。

参考文献

- [1] 任菊香, 刘忠宝. 融合多粒度语义特征的中文情感分析方法[J]. 华东师范大学学报(自然科学版), 2023(6): 95-107.
- [2] 王昆, 郑毅, 方书雅, 等. 基于文本筛选和改进 BERT 的长文本方面级情感分析[J]. 计算机应用, 2020, 40(10): 2838-2844.
- [3] Poria, S., Cambria, E., Hazarika, D., Majumder, N., Zadeh, A. and Morency, L. (2017) Context-Dependent Sentiment Analysis in User-Generated Videos. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Vancouver, July 2017, 873-883. <https://doi.org/10.18653/v1/p17-1081>
- [4] Majumder, N., Poria, S., Hazarika, D., Mihalcea, R., Gelbukh, A. and Cambria, E. (2019) DialogueRNN: An Attentive RNN for Emotion Detection in Conversations. *Proceedings of the AAAI Conference on Artificial Intelligence*, **33**, 6818-6825. <https://doi.org/10.1609/aaai.v33i01.33016818>
- [5] Poria, S., Majumder, N., Mihalcea, R. and Hovy, E. (2019) Emotion Recognition in Conversation: Research Challenges, Datasets, and Recent Advances. *IEEE Access*, **7**, 100943-100953. <https://doi.org/10.1109/access.2019.2929050>
- [6] Liu, Y., Bangerter, L.R., Rovine, M.J., Zarit, S.H. and Almeida, D.M. (2018) Intrinsic Emotional Fluctuation in Daily Negative Affect across Adulthood. *The Journals of Gerontology: Series B*, **73**, 100-112.
- [7] Lee, C., Busso, C., Lee, S. and Narayanan, S.S. (2009) Modeling Mutual Influence of Interlocutor Emotion States in Dyadic Spoken Interactions. *Inter Speech 2009*, Brighton, 6-10 September 2009, 1983-1986. <https://doi.org/10.21437/interpeech.2009-480>
- [8] Tang, D., Qin, B. and Liu, T. (2015) Document Modeling with Gated Recurrent Neural Network for Sentiment Classification. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, Lisbon, 17-21 September 2015, 1422-1432. <https://doi.org/10.18653/v1/d15-1167>
- [9] Naseem, U., Razzak, I., Musial, K. and Imran, M. (2020) Transformer Based Deep Intelligent Contextual Embedding for Twitter Sentiment Analysis. *Future Generation Computer Systems*, **113**, 58-69. <https://doi.org/10.1016/j.future.2020.06.050>
- [10] Sun, G., Zhang, C. and Woodland, P.C. (2021) Transformer Language Models with LSTM-Based Cross-Utterance Information Representation. *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, 6-11 June 2021, 7363-7367. <https://doi.org/10.1109/icassp39728.2021.9414477>
- [11] Song, K., Bing, L., Gao, W., Lin, J., Zhao, L., Wang, J., et al. (2019) Using Customer Service Dialogues for Satisfaction Analysis with Context-Assisted Multiple Instance Learning. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, Hong Kong, November 2019, 198-207. <https://doi.org/10.18653/v1/d19-1019>
- [12] Mao, Y., Cai, F., Guo, Y. and Chen, H. (2022) Incorporating Emotion for Response Generation in Multi-Turn Dialogues. *Applied Intelligence*, **52**, 7218-7229.
- [13] Devlin, J. (2018) Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding.
- [14] Koroteev, M.V. (2021) BERT: A Review of Applications in Natural Language Processing and Understanding.
- [15] Munikar, M., Shakya, S. and Shrestha, A. (2019) Fine-Grained Sentiment Classification Using Bert. *2019 Artificial Intelligence for Transforming Business and Society (AITB)*, Kathmandu, 5 November 2019, 1-5. <https://doi.org/10.1109/aitb48515.2019.8947435>
- [16] Zhang, Z., Guo, T. and Chen, M. (2021) Dialogue BERT. *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, New York, 1-5 November 2021, 3647-3651. <https://doi.org/10.1145/3459637.3482085>
- [17] Yang, H. and Shen, J. (2021) Emotion Dynamics Modeling via BERT. *2021 International Joint Conference on Neural Networks (IJCNN)*, Shenzhen, 18-22 July 2021, 1-8. <https://doi.org/10.1109/ijcnn52387.2021.9533860>
- [18] Ni, J., Young, T., Pandelea, V., Xue, F. and Cambria, E. (2022) Recent Advances in Deep Learning Based Dialogue Systems: A Systematic Survey. *Artificial Intelligence Review*, **56**, 3055-3155. <https://doi.org/10.1007/s10462-022-10248-8>
- [19] SenticNet. <https://sentic.net/>
- [20] Poria, S., Hazarika, D., Majumder, N., Naik, G., Cambria, E. and Mihalcea, R. (2019) MELD: A Multimodal Multi-Party

Dataset for Emotion Recognition in Conversations.

- [21] GitHub (2018) JDDC-Baseline-Seq2Seq. <https://github.com/SimonJYang/JDDC-Baseline-Seq2Seq/tree/master/data>
- [22] Kim, T. and Vossen, P. (2021) Emoberta: Speaker-Aware Emotion Recognition in Conversation with Roberta.
- [23] Kumar, A., Irsoy, O., Ondruska, P., Iyyer, M., Bradbury, J., Gulrajani, I. and Socher, R. (2016) Ask Me Anything: Dynamic Memory Networks for Natural Language Processing. *Proceedings of the 33rd International Conference on Machine Learning*, New York, 19-24 June 2016, 1378-1387.
- [24] Li, J., Zhang, M., Ji, D. and Liu, Y. (2020) Multi-Task Learning with Auxiliary Speaker Identification for Conversational Emotion Recognition.
- [25] Ghosal, D., Majumder, N., Poria, S., Chhaya, N. and Gelbukh, A. (2019) Dialoguecn: A Graph Convolutional Neural Network for Emotion Recognition in Conversation. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, Hong Kong, November 2019, 154-164. <https://doi.org/10.18653/v1/d19-1015>