

基于人工与 ChatGPT 标注的推文情感分析对比研究

杨 艺¹, 黄镜月¹, 贺品尧¹, 荣 婷²

1. 重庆工商大学 人工智能学院, 重庆 400067

2. 重庆工商大学 重庆高校网络舆情与思想动态研究咨政中心, 重庆 400067

摘 要:目的 针对特定推文情感分析任务中标注数据的困难和由于标注不准确带来的分类结果不尽如人意问题,提出一种机器标注数据的方法来研究深度学习模型对人工标注和机器标注推文数据情感分类的性能表现差异。方法 研究中,对于统一的标签体系,分别对推文数据进行人工标注和运用 ChatGPT 模型接口标注,再采用 BERT-TextCNN 深度学习混合模型,对经过人工标注和 ChatGPT 标注的数据集进行情感分类。结果 实验结果表明:人工标注数据集在整体性能上表现出更高的准确性和可信度,但是在某些推文数据上,ChatGPT 大模型以其比人更丰富的知识储备,可以生成比人更客观科学的可解释性标注,在情感分类结果上呈现出一定的优势,人工标注和机器标注方法各具优劣;由此可以得出对于文本情感分类任务,机器标注是一种可行的标注方法。结论 在实际应用场景中,可以根据任务需求灵活选择和结合两种标注方法,充分利用两者之间的优势,以达到更佳的分析性能和效果。

关键词:人工标注; ChatGPT 标注; 推文; 情感分析; BERT-TextCNN

中图分类号: TP391.1 文献标识码: A doi:10.16055/j.issn.1672-058X.2025.0004.012

Comparative Study on Sentiment Analysis of Tweets Based on Manual and ChatGPT Annotation

YANG Yi¹, HUANG Jingyue¹, HE Pinyao¹, RONG Ting²

1. School of Artificial Intelligence, Chongqing Technology and Business University, Chongqing 400067, China

2. Chongqing Center for Research and Consultancy of Cyber Public Opinion and Ideological Development in Universities, Chongqing Technology and Business University, Chongqing 400067, China

Abstract: Objective This study addresses the challenges of annotating data for specific tweet sentiment analysis tasks and the issues arising from inaccurate annotations leading to unsatisfactory classification results. A method for machine annotation of data was proposed to investigate the performance differences in sentiment classification of tweets annotated by human annotators and the ChatGPT model. **Methods** In this study, for a unified labeling system, tweet data was annotated both manually and using the ChatGPT model interface, followed by sentiment classification using a BERT-TextCNN hybrid deep learning model on both datasets. **Results** Experimental results indicated that the manually annotated dataset exhibited higher overall accuracy and reliability. However, for certain tweet data, the ChatGPT model, with its richer knowledge base, can produce more objective and scientifically interpretable annotations, showing certain advantages in sentiment classification results. Both human and machine annotation methods have their strengths and weaknesses. Therefore, it can be concluded that machine annotation is a feasible labeling method for text sentiment classification tasks. **Conclusion** In practical applications, it is advisable to flexibly choose and combine both annotation methods based on task requirements and fully leverage the strengths of these two methods to achieve better analytical performance and outcomes.

Keywords: manual annotation; ChatGPT annotation; tweets; sentiment analysis; BERT-TextCNN

收稿日期: 2023-12-17 修回日期: 2024-04-11 文章编号: 1672-058X(2025)04-0095-07

基金项目: 国家社会科学基金青年项目(22CZZ051)。

作者简介: 杨艺(1971—), 女, 四川成都人, 教授, CCF 会员, 从事大数据分析处理, 信息系统开发研究。Email: yangchongyi@ctbu.edu.cn.

引用格式: 杨艺, 黄镜月, 贺品尧, 等. 基于人工与 ChatGPT 标注的推文情感分析对比研究[J]. 重庆工商大学学报(自然科学版), 2025, 42(4): 95-101.

YANG Yi, HUANG Jingyue, HE Pinyao, et al. Comparative study on sentiment analysis of tweets based on manual and ChatGPT annotation[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2025, 42(4): 95-101.

1 引言

情感分析是自然语言处理中的重要方面,旨在确定给定文本的情感倾向,如正面、负面或中性。在许多实际应用中,利用文本情感分析(TEA)理解和解释文本中的情感信息非常有用。Twitter 平台可以让用户更新不超过 140 个字符的消息,用户可以用几句话表达自己的感想或对某主题进行评论,故其推文具有短小精悍的特点,从平台上可以容易地获得大规模的数据。通过对其推文进行情感分析,可以获得关于用户情感倾向的洞察,从而支持决策制定、了解用户偏好,收集舆论走向等。

传统使用的模型有基于词袋模型的朴素贝叶斯分类器、支持向量机(SVM)以及基于规则的方法等^[1]。然而,这些传统方法在处理如推文这样的短文本数据时存在一些局限性,它们无法很好地捕捉文本中的上下文信息和语义关系,导致性能不尽如人意。此外,人工标注可能受到标注人主观因素的影响,从而导致标注结果可能不够准确和一致,直接影响模型对其的分析。

ChatGPT 的出现为短文本情感分析带来了新的可能性。ChatGPT 能够对数据进行可解释性标注,为研究人员提供了一种全新的视角。当前,关于机器对短文本标注的研究正在不断增加。研究人员提出了对原始数据采用人工和机器标注相结合的方法来研究短文本情感分析。这种方法在提高标注效率和标注质量的同时,也能减少主观误差,从而有效地提高情感分析的准确性。这一新的研究方向为短文本情感分析领域带来了更为全面和深刻的理解,为未来的研究提供了更为广阔的发展空间。

2 相关工作

2.1 情感分析

自从 Bo Pang 2002 年提出情感分析,就一直受到广泛的关注,特别是在在线评论的情感倾向性分析方面取得了显著的发展。情感分析是研究人们对于产品、服务、个人、事件、主题及其属性等实体对象所持有的意见、情感、评价、看法和态度等主观感受的研究领域。情感分析的研究对象可以是任何可以传达人类情感的事物,包括文本、图片、声音和视频等^[2]。Twitter 作为社交媒体平台,其推文数据非常丰富,通过对这些数据进行情感分析,可以发现和利用其中高价值的隐含信息,并在一定程度上揭示信息之间的关联^[3]。本文首先爬取了 Twitter 上与 XXX 战争相关的推文评论数据,然后对其进行清洗和标注,再通过情感分析模型判断其情感倾向。

2.2 推文评论数据情感倾向标签

传统的人工标注数据集因标注者之间的差异,可能导致标签偏见或主观判断的存在,而机器学习方法可以借助预定义的规则或模型对数据进行自动标注,但这样做可能会受到规则设计或模型选择的限制,导致标注结果不够准确或完善^[4]。使用深度神经网络模型进行数据标注也是一种方法,但在没有足够大规模、高质量数据进行训练的情况下,深度学习模型可能会过度拟合或产生不稳定的标注结果^[5]。由于 ChatGPT 模型是基于大规模语料库进行预训练的,它可以从海量的数据中学习并捕捉文本的语义和上下文信息,从而能够在一定程度上避免人工标注受限于标注者的知识面对标签的解释能力等主观原因带来的标注差异影响,提供更一致的标注结果,同时减少人工标注数据集的时间和劳力成本。

3 方法

首先,对经过预处理的 XXX 战争相关推文评论数据,用两种方法进行标注,得到一个由人工标注的数据集和另一个由 ChatGPT 生成的标注数据集。接着,采用了 BERT、TextCNN 和 BERT-TextCNN 模型进行情感分析,以评估不同标注数据在模型性能和效果上的差异。

3.1 基于 ChatGPT 可解释性标签处理

3.1.1 标签体系

本研究的目标即是通过分析 Twitter 用户对于 XXX 战争这个主题的评论,了解其态度、看法和情感表达,为进一步的研究和决策提供有价值的信息。根据目前的标签体系准则以及研究目标构建了标签体系,将评论数据的情感倾向划分为以下 5 个类别:支持 A 方、支持 B 方、反对 A 方、反对 B 方以及倾向不明确,并用 0-4 的数值与之对应,具体如表 1 所示。为了确保标签的准确性和一致性,在进行标注过程时采取了严格的标注准则,并尽可能避免主观偏见。同时,还进行了标签之间的相互对比,以确保各个情感类别之间的区分度和明确性。这样做有助于提高标注质量,并为后续的情感分析和研究提供可靠的数据基础。

表 1 XXX 战争评论数据标签体系

Table 1 Labeling system for XXX war commentary data

标签值	情感分类
0	倾向不明确
1	支持 A 方
2	支持 B 方
3	反对 A 方
4	反对 B 方

3.1.2 ChatGPT 标注

目前,数据标注主要依靠人力完成。然而,当标注员面临复杂的标注任务或者庞大的标注数据量时,常常会面临巨大的心理压力。同时,数据标注工作本身具有高度的重复性,标注时间紧迫,并且缺乏严格的质量审核流程,这将会导致标注任务的合格率低、标注不完备等问题,不准确或不完整的标注数据可能导致训练出的模型产生误导性的结果。此外,人工标注标签还存在主观性的问题,即使多名专家参与标注,由于主观性的差异,标签的一致性可能受到影响,同样会影响后续模型分析的准确性。同时,标注不及时也会导致在实时应用中无法及时获得准确的数据分析和决策支持^[6]。为解决这些问题,深度学习方法被引入用于自动标注文本情感,然而,现有的深度学习模型在自动标注方面仍存在一些局限。首先,小模型的复杂性和参数数量有限,难以准确捕捉文本中的细微情感和语义。其次,这些小模型往往缺乏可解释性,难以提供对标注结果的解释和理解,使其在实际应用中的可靠性和可行性受到质疑。为了克服这些局限,考虑利用当前的大模型 ChatGPT 来进行文本数据的自动标注,并提供可解释性的标注结果。ChatGPT 作为一种强大的语言模型,具有广泛的语义理解和生成能力。通过训练 ChatGPT 模型进行文本情感分析的自动标注,可以有效利用其强大的语言处理能力,从而提高标注的准确度和效率。利用 ChatGPT 对文本进行标注的主要处理流程如图 1 所示。

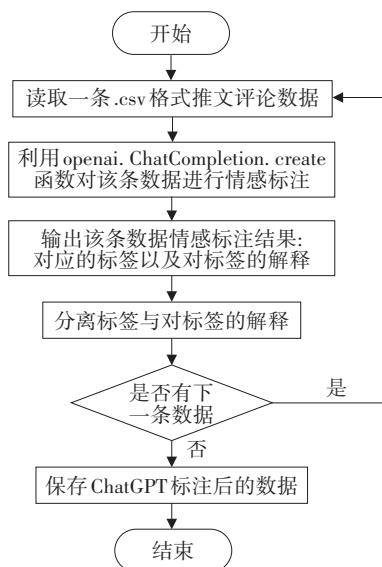


图 1 ChatGPT 对文本进行标注的处理流程

Fig. 1 Processing flow of ChatGPT for text annotation

3.2 深度学习模型

3.2.1 BERT (Bidirectional Encoder Representations Form Transformers)

BERT 模型在自然语言处理领域显著提升了处理结果的准确度,BERT 的网络结构是基于 Transformer 的架构^[7]。

BERT 的输入层和输出层都是矩阵形式,因此可以串联 N 个 BERT 层。它的输入层通过分词、特殊令牌添加、位置嵌入和词向量映射等处理步骤,将原始文本转换为一个包含词向量表示的输入序列^[8]。在这些步骤中,位置嵌入的使用对于模型理解句子的含义具有重要影响,它通过采用正弦波函数来编码令牌的位置信息,类似于模拟信号的周期性变化,位置嵌入的计算方法如下:

$$P_E(p_{os}, 2i) = \sin\left(\frac{p_{os}}{10\,000^{\frac{2i}{d_{model}}}}\right)$$

$$P_E(p_{os}, 2i+1) = \cos\left(\frac{p_{os}}{10\,000^{\frac{2i}{d_{model}}}}\right) \quad (1)$$

式(1)中, $P_E(p_{os}, i)$ 表示位置 p_{os} 和维度 i 处的位置编码值,其中, p_{os} 为位置信息, i 为维度的索引, d_{model} 表示词向量的维度。在计算位置嵌入时,位置嵌入向量的每个维度都根据位置 p_{os} 和维度索引 i 使用不同的计算公式进行计算。

此外,BERT 模型通过掩码语言模型实现了真正意义上的双向编码。它随机地屏蔽输入文本中大约 15% 的令牌,并通过被屏蔽的令牌位置的最终隐藏层向量传递给分类器来预测被屏蔽的令牌。确定屏蔽的单词后,并没有直接删除它们,而是随机替换约 80% 的被屏蔽令牌,再用任意单词替换余下的 10%,最后只保留 10% 的原始令牌。这样的设计旨在增强模型的鲁棒性,避免出现模型无法识别的单词以提高其泛化能力^[9]。

3.2.2 Text CNN(Text Convolutional Neural Network)

2014 年,Yoon Kim 对 CNN 的输入层进行了一些修改,提出了文本分类模型 Text CNN^[10]。Text CNN 除了保留原始 CNN 的特征提取能力外,还增强了对文本特征的提取能力。在图 2 所示的 Text CNN 模型中,设置了不同大小的过滤器,通过卷积层和池化层对词向量形成的矩阵进行处理,卷积核滑过矩阵的一行,也就是一个单词,卷积核和矩阵的宽度相同,而高度可以根据任务需要进行设置。通过卷积操作,可以提取 N 个相邻单词的特征。最后使用 Softmax 函数对输出进行分类^[11]。

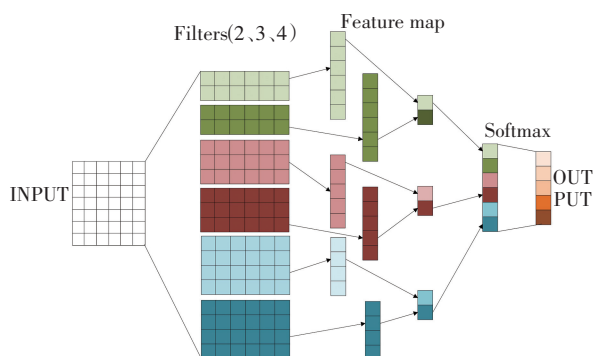


图 2 Text CNN 模型

Fig. 2 Text CNN model

卷积运算可以用以下公式表示:

$$Y_i = W \cdot A[i:i+H-1], i=1, 2, \dots, S-H+1 \quad (2)$$

则最终提取的特征 F_i 可表示为

$$F_i = f(Y_i + B) \quad (3)$$

其中, A 表示输入的短文本矩阵, S 表示短文本中的词数, W 是卷积核的权重参数, Y_i 代表卷积运算得到的特征图中的第 i 个特征, B 是偏置项, 用于调整特征图中每个特征的偏移量, f 是激活函数, F_i 是经过激活函数后得到的最终特征。通过这样的卷积运算, 可以一次性提取出 H 个特征, 从而丰富短文本的表示能力。经过卷积运算后, 将短文本中的局部情感信息传递到池化层, 压缩提取主要局部特征, 去除冗余信息。然后将得到的局部特征拼接构成新的矩阵, 进一步减少参数数量、数据维数以及避免过拟合^[12]。

3.2.3 BERT-Text CNN

在本研究中, 采用了 BERT-Text CNN 的混合模型, 它结合了 BERT 和 Text CNN 的优势。首先, 使用 BERT 预训练语言模型从文本中提取语义表示和上下文信息。BERT 能够有效地捕获词之间的语义关系, 为后续的情感分析任务提供丰富的特征表示。然后, 特征提取模块将从 BERT 模型接收到的文本特征表示输入到 Text CNN 中, 以从文本中提取局部特征。Text CNN 模型由多个卷积层组成, 每个卷积层滑动固定大小的卷积核在文本上进行特征提取, 有效地捕获局部特征。接着, 池化层用于集成每个卷积核提取的特征, 从而减少特征的维度并保留最重要的信息。这样, 就能够提取并集成单词之间的局部依赖关系, 将这些特征用于情感分析任务, 让模型更好地理解文本中的语义信息, 并有效地进行情感倾向分类^[13-14]。通过 BERT-Text CNN 的结合, 可以充分利用 BERT 强大的上下文理解能力和 Text CNN 的局部特征提取能力, 提高情感分析任务的性能和准确性。BERT-Text CNN 结构如图 3 所示。

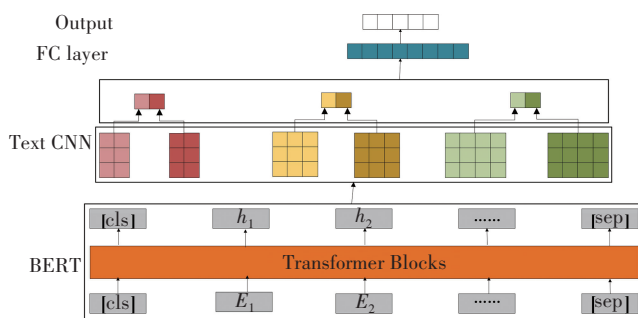


图 3 BERT-Text CNN 结构

Fig. 3 BERT-Text CNN hybrid architecture

4 仿真实验与结果分析

4.1 实验数据标注

实验采用的数据共包含 10 000 多条关于 XXX 战争的推文评论, 为了确保数据的质量, 进行了预处理操作, 包括去除无意义和重复的数据。在最终的处理过程中, 剩余 9 450 条数据, 分别进行了人工标注和 ChatGPT 标注, 得到 5 个情感分类的实验数据集, 具体情况如表 2 所示。

表 2 实验数据集标注结果分布

Table 2 Distribution of annotation results in experimental dataset

标签值	情感分类	人工标注数 数据集/条	ChatGPT 标注 数据集/条
0	倾向不明确	5 982	4 058
1	支持 A 方	701	2 009
2	支持 B 方	1 336	1 494
3	反对 A 方	1 210	1 608
4	反对 B 方	221	281

为确保研究结果的可靠性和一致性, 邀请了 5 位专家参与推文评论的情感分类标注任务, 每位专家都独立完成标注工作。在开始标注任务之前, 将推文评论数据以 Excel 表格的形式交付给参与标注的专家团队, 并在每行标示了一个唯一的标识符, 以便专家进行标注。在交付数据之前, 向专家团队提供了详细的标注准则或指南。这些准则包含了情感类别的标签定义和解释, 旨在确保专家在进行标注时能够准确理解情感类别的含义。为了确保标注的一致性和准确性, 在实际标注之前为参与任务的每位专家提供了培训。培训的目的是帮助专家们深入了解标注任务的目标和标准, 并降低标注中的主观性因素, 从而提高标注的一致性。每位专家在独立标注过程中, 根据个人理解和判断对推文评论进行情感分类, 从而形成了独立的标注结果。采用独立标注的方式有助于避免专家之间的互

相影响和偏见,从而保持标注结果的客观性。完成独立标注后,专家团队进行了一致性检查,以评估标注结果之间的一致性水平。如果发现标注结果存在较大差异或不一致,专家团队会进行讨论。在讨论过程中,共同审查标注准则,澄清疑惑,并努力达成一致意见。讨论的过程可能导致标注准则的更新或进行进一步的培训,以确保标注结果更加一致和准确。通过这样的反复讨论和调整,最终可以达成一致的标注结果,确保了标注任务的可靠性和准确性。

为了让 ChatGPT 对推文评论进行自动标注,编写了 Python 程序。首先要配置网络环境,确保自动标注程序能持续连接在运行 OpenAI GPT 3.5 访问接口的服务器上。其次,将需要访问 OpenAI GPT 3.5 接口的数据组织成一个 Excel 文档,并在特定的列中存放原始推文评论数据。随后,通过 Pandas 库按行读取 Excel 文件中数据,调用 OpenAI 库以对话方式让 ChatGPT 对该行数据进行标注。在 OpenAI 库中最核心的是使用了 OpenAI GPT-3.5 API 的一个函数 ChatCompletion,调用其 Create 方法可以创建一个对话模型的请求,以便与 GPT-3.5 模型进行交互,生成自然语言文本的回答。其中的参数需按照 OpenAI 官网提供的格式进行编写,具体如表 3 所示。为了增强程序的健壮性,需捕获和处理可能出现的异常情况。在 GPT-3.5 模型访问数据时,为减少访问的时间,可以根据设备性能选择合适的线程数,充分利用多线程进行并发访问,加快数据的获取和处理过程。最后,ChatCompletion 函数的 Message 参数对于 ChatGPT 的回答至关重要,一定根据自己的具体需求设置好 Message 参数的内容,即准确的提问词,才能确保获取更准确地回答。

表 3 ChatGPT 参数设置
Table 3 ChatGPT parameter settings

参 数	说 明	取 值
Model	用于指定调用的模型	GPT 3.5
Message	用于说明采用对话模式,并与传入的推文评论数据进行绑定	根据输入的判断规则 0-4,对以下文本进行标注以及标注解释
Max_tokens	用于限制 OpenAI 生成的回答长度	100
Temperature	用于控制回答的随机性,值越大回答越具有随机性	1.0

4.2 实验设置

本实验中将 BERT、Text CNN 和 BERT-Text CNN

深度学习模型应用于人工标注和 ChatGPT 标注数据集的情感分析,并进行对比评估,以探索不同标注方法在特定推文情感分析任务中的性能表现。

在实验过程中,将数据集的 70% 用作训练集,而余下的 30% 则被用作测试集,为了防止模型过度复杂化而导致的过拟合问题,对超参数进行了深入研究,以寻求最佳的模型分类性能。在优化阶段,采用了 Adam 优化器,并结合了分类交叉熵损失函数对模型进行了训练。分类交叉熵损失函数使得模型能够分配独立的标签概率,这对于解决多标签分类问题是必不可少的要素。在表 4 中,详细列出了通过多次实验获得的最优超参数配置^[15-16]。

表 4 BERT-TextCNN 模型的超参数
Table 4 Hyperparameters of the BERT-TextCNN model

Hyper-parameters	Values
Learning rate	0.001
Loss Function	CrossEntropyLoss
Optimizer	Adam
Batch size	16
Dropout	0.2
Convolutional size	9
Kernel sizes	[2,3,4]
Epochs	10

4.3 实验结果及分析

两种标注的数据集在 3 种模型运行的实验结果如表 5 所示。与 ChatGPT 标注的数据集相比,人工标注的数据集在实验结果中都表现得相对较高。而 3 种模型中,BERT-Text CNN 模型实验结果最好,因为 BERT-Text CNN 模型能够更好地利用语义信息和上下文关系,从而对情感分类任务有更准确的理解和处理能力,具体如图 4 所示。

表 5 ChatGPT 标注的数据集的实验结果

Table 5 Experimental results on the ChatGPT annotated dataset

标注类型	模 型	Accurary	Precision	Recall	F1 score
ChatGPT	Text CNN	0.590 3	0.574 6	0.590 3	0.530 1
	BERT	0.672 9	0.691 5	0.672 9	0.673 8
	BERT-TextCNN	0.850 5	0.850 5	0.850 5	0.850 5
	Text CNN	0.866 0	0.884 0	0.866 0	0.803 9
人工	BERT	0.932 1	0.872 1	0.932 1	0.866 5
	BERT-TextCNN	0.975 1	0.975 1	0.975 1	0.975 1
	TextCNN				

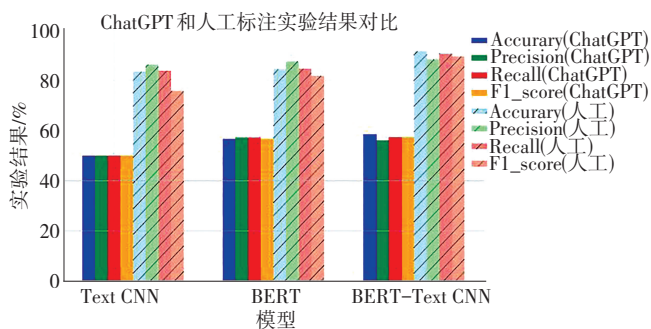


图 4 模型结果对比图

Fig. 4 Comparison of results between two models

在本次实验中,发现 ChatGPT 标注数据集的各项指标相较于人工标注数据集,其性能表现差一些,这一结果可能受到数据集存在数据分类不平衡以及个别类别数据较少的影响。在实验结果不理想时,标注专家共同对 ChatGPT 标注的数据集进行了深入的研究和分析。在这个过程中,观察到在部分数据中,存在 ChatGPT 对文本的标注表现更为准确的情况,比如表 6 给出的示例。这可以归因于 ChatGPT 在预训练过程中接触到了大量的背景知识,并具备较强的文本理解能力,使得它在特定任务、领域或含有特定背景信息的文本上表现出一定的优势,从而提高了情感分类的准确性。

同时,也意识到 ChatGPT 在情感标注中存在一些问题。其中,噪声问题是一个主要因素,由于 ChatGPT 预训练时使用了大量非结构化文本数据,其中可能含有噪声和错误信息,这可能导致标注过程中产生误差。

表 6 ChatGPT 标注更准确的示例

Table 6 Examples of more accurate annotations by ChatGPT

推文评论	ChatGPT 的解释	ChatGPT 标签	人工标签
Welcome to the day of the invasion, to invasion day!	The sentence sarcastically refers to the day of the invasion, implying that the speaker is against Ukraine and views the invasion as a negative event.	4	0

此外,ChatGPT 的上下文理解也存在一定的局限,虽然预训练过程中学习了丰富的语言知识,但在特定标注任务上可能表现较差。其次,由于 ChatGPT 是一个大模型,在标注数据集较小的情况下可能存在过拟合,导致在测试集或实际应用中性能下降。同时,ChatGPT 预训练数据的多样性也有待思考,它可能未涵

盖特定标注任务所需的所有领域和文本类型,从而影响其对特定任务的标注准确性和可靠性。在表 7 中展示了一个 ChatGPT 对数据进行过度解读的示例。

表 7 ChatGPT 对数据进行过度解读的示例

Table 7 Examples of overinterpretation by ChatGPT on data

推文评论	ChatGPT 的解释	ChatGPT 标签	人工标签
This whole Russia/Ukraine and Ukraine to the thing is feeling a lot like conflict between Germany the whole Germany/Poland and Poland in 1939, thing circa 1939, except implying that they see this time there won't be Russia's actions as any potential ally on a aggressive and potentially potential Eastern front that leading to a larger we can count on when we conflict. The use of the finally pull our heads out hashtag #UkraineInvasion of the sand.	The speaker compares the situation between Russia and Ukraine to the conflict between Germany and Poland in 1939, implying that they see Russia's actions as aggressive and potentially leading to a larger conflict. The use of the hashtag #UkraineInvasion further emphasizes their disapproval of Russia's actions in Ukraine.	3	0

从以上分析可知,采用 ChatGPT 对短文本进行标注是可行的,当标注专家感觉标注困难时,可以考虑采用 ChatGPT 的标注结果进行辅助,同时,可以提升标注的效率。

5 结 论

在短文本情感分析任务中,尽管人工标注数据集展现出较高的性能,具备更好的准确性和可信度,我们也留意到 ChatGPT 标注数据集在某些特定任务上呈现出独特的优势,为短文本情感分析领域引入了一种创新的标注方法。未来的研究将深入探讨并改进 ChatGPT 标注数据集的性能。通过探索优化方法,例如增加多样性的预训练数据和提升模型的语义理解能力,以使 ChatGPT 在情感分类任务中发挥更大的潜力。同时,今后将着眼于如何巧妙地利用人工标注和 ChatGPT 标注之间的互补性,以进一步提高短文本情感分析的性能和效果,实现更为全面的情感分析。这不仅将有助于改善 ChatGPT 标注数据集在短文本情感分析中的表现,还其他自然语言处理任务提供更为精确的标注数据。

此外,对于 ChatGPT 在情感标注任务中整体性能

相对较低的问题,需要深入认识其局限性,并努力最大化其优点。未来的研究可以集中于改进 ChatGPT 的上下文理解,解决噪声问题,提高对特定领域语境的适应能力,以及将 ChatGPT 与其他模型进行融合等方面,为未来的研究提供更为综合且精准的情感标注结果,从而推动情感分析领域的发展。

参考文献(References):

- [1] PENG S, CAO L, ZHOU Y, et al. A survey on deep learning for textual emotion analysis in social networks[J]. Digital Communications and Networks, 2022, 8(5): 745–762.
- [2] HE Y X, SUN S T, NIU F F, et al. A deep learning model with enhanced sentiment semantics for Weibo sentiment analysis[J]. Journal of Computer Science and Technology, 2017, 40(4): 773–790.
- [3] WANG L X. A review of preprocessing and learning research on Weibo short texts[J]. Library and Information Service, 2013, 57(11): 125.
- [4] SNOW R, O'CONNOR B, JURAFSKY D, et al. Cheap and fast: But is it good: Evaluating non-expert annotations for natural language tasks[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing-EMNLP'08. Morristown, NJ, USA: ACL, 2008: 254–263.
- [5] DOS SANTOS C N, GATTI M. Deep convolutional neural networks for sentiment analysis of short texts[J]. COLING 2014-25th International Conference on Computational Linguistics, Proceedings of COLING 2014: Technical Papers, 2014: 69–78.
- [6] 蔡莉,王淑婷,刘俊晖,等.数据标注研究综述[J].软件学报,2020,31(2): 302–320.
CAI Li, WANG Shu-ting, LIU Jun-hui, et al. Survey of data annotation[J]. Journal of Software, 2020, 31(2): 302–320.
- [7] 史振杰,董兆伟,庞超逸,等.基于BERT-CNN的电商评论情感分析[J].智能计算机与应用,2020,10(2): 7–11.
SHI Zhen-jie, DONG Zhao-wei, PANG Chao-yi, et al. Sentiment analysis of e-commerce reviews based on BERT-CNN[J]. Intelligent Computer and Applications, 2020, 10(2): 7–11.
- [8] ZHANG T, YOU F. Research on short text classification based on TextCNN[J]. Journal of Physics: Conference Series, 2021, 1757(1): 012092.
- [9] DONG J, HE F, GUO Y, et al. A commodity review sentiment analysis based on BERT-CNN model[C]//Proceedings of the 5th International Conference on Computer and Communication Systems. Piscataway: IEEE Press, 2020: 143–147.
- [10] SHAN D, LI H. Group emotion recognition for Weibo topics based on BERT with TextCNN[J]. American Journal of Information Science and Technology, 2023: 95–100.
- [11] ABAS A R, ELHENAWY I, ZIDAN M, et al. BERT-CNN: a deep learning model for detecting emotions from text[J]. Computers, Materials & Continua, 2022, 71(2): 2943–2961.
- [12] 宋鹏,耿成,李泽.基于卷积神经网络的文本分类研究[C]//2019 计算机网络、电子与自动化国际会议(ICCNEA). IEEE, 2019: 229–232.
SONG Peng, GENG Cheng, LI Ze. Research on Text Classification Based on Convolutional Neural Network[C]//2019 International Conference on Computer Networks, Electronics and Automation (ICCNEA). IEEE, 2019: 229–232.
- [13] DONG M, LI Y, TANG X, et al. Variable convolution and pooling convolutional neural network for text sentiment classification[J]. IEEE Access, 2020, 8: 16174–16186.
- [14] 王煜涵,张春云,赵宝林,等.卷积神经网络下的Twitter文本情感分析[J].数据采集与处理,2018,33(5): 921–927.
WANG Yu-han, ZHANG Chun-yun, ZHAO Bao-lin, et al. Sentiment analysis of Twitter data based on CNN[J]. Journal of Data Acquisition and Processing, 2018, 33(5): 921–927.
- [15] ZHAO J, GUI X, ZHANG X. Deep convolution neural networks for Twitter sentiment analysis[J]. IEEE Access, 2018, 6: 23253–23260.
- [16] MAN R, LIN K. Sentiment analysis algorithm based on BERT and convolutional neural network[C]//Proceedings of the IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers. Piscataway: IEEE Press, 2021: 769–772.

责任编辑:陈芳