

李奕霏, 陈珊. 短视频传播受众情绪预判研究[J]. 智能计算机与应用, 2024, 14(10): 87-91. DOI: 10.20169/j.issn.2095-2163.241011

短视频传播受众情绪预判研究

李奕霏, 陈珊

(临沂大学传媒学院, 山东临沂 276000)

摘要: 在短视频传播中, 用户情感分析是探索视频媒体传播规律的基础。短视频弹幕包含了分析用户情感倾向的重要信息。因此, 通过短视频弹幕识别受众的情感倾向对视频媒体传播具有重要意义。本文利用视频媒体中弹幕信息来分析传播受众的情绪态势, 以计算机深度学习模型为核心设计一个新型的面向短视频弹幕情绪识别的预测模型, 通过将短视频传播受众的情绪分成5类, 构建用户情绪空间, 并借助浅度学习模型来实时预判受众群体的情绪倾向。

关键词: 短视频; 弹幕; 情绪预判; 预测模型

中图分类号: TP391.1

文献标志码: A

文章编号: 2095-2163(2024)10-0087-05

Research on predicting the audience emotions in short video dissemination

LI Yifei, CHEN Shan

(School of Media, Linyi University, Linyi 276000, Shandong, China)

Abstract: In short-video dissemination, user emotion analysis is the foundation for exploring the dissemination laws of video media. Short-video bullet comments contain important information for analyzing the emotional tendencies of users. Therefore, identifying the emotional tendencies of the audience through short-video bullet comments is of great significance for video media dissemination. This paper uses the bullet comment information in video media to analyze the emotional situation of the dissemination audience. And the paper designs a new prediction model for short-video bullet comment emotion recognition centered around a computer deep learning model. By dividing the emotions of the short-video dissemination audience into five categories, a user emotion space is constructed, and a shallow learning model is used to predict the emotional tendencies of the audience group in real time.

Key words: short videos; bullet comments; emotion prediction in advance; prediction model

0 引言

情绪是对一系列主观认知经验的通称, 是人对外观事物的态度体验以及相应的行为反应。受众情绪是指在特定时间段内受众对待新闻宣传事件表现出的趋同性的情绪体验。其凝结了公众记忆, 反映着社会共同心态, 表达了集体诉求, 是受众基本的价值判断和品牌传播行为碰撞形成的集体情感共鸣。传播受众的情绪感知与短视频传播力密切相关。受众情绪感知为正面时, 随着传播力的增强, 短视频宣传会获得更多的情绪共鸣和价值认可。受众情绪感知为负面时, 随着传播力的增强, 产品短视频的宣传会激发受众更为恶劣的情绪, 甚至感受最终演变为舆情。因此动态实时感知短视频或者直播视频受众

的情绪变化是提升传播力的基础支撑。本文利用视频媒体中弹幕信息来分析传播受众的情绪态势, 以计算机深度学习模型为核心构建短视频受众情绪预测模型, 来实时预判受众群体的情绪倾向。

在此设计一个新型的面向短视频弹幕情绪识别的预测模型。首先将短视频传播受众的情绪分成5类, 构建用户情绪空间; 再通过浅度学习模型预处理获取视频弹幕, 把预处理后的文本向量输入到深度模型中; 由浅度学习模型完成弹幕情绪识别。该预测模型不仅充分提取弹幕上下文的文本和语义特征, 而且提高模型的可解释性。量化分析结果表明, 分类精度达到87.7%, 在训练集和测试集上的准确率达到99.6%, 因此在短视频弹幕情绪识别中该预测模型具有较高的准确率, 可作为短视频传播机

基金项目: 2023年度山东省社会科学规划研究项目(22CLYJ04)。

作者简介: 李奕霏(1991-), 女, 博士, 讲师, 主要研究方向: 新媒体传播, 大数据, Email: liyifei@lyu.edu.cn; 陈珊(1978-), 女, 博士, 教授, 主要研究方向: 自然语言处理, 舆情分析, 情感计算等。

收稿日期: 2024-02-20

制分析的基础工具。

1 短视频弹幕概述

随着人们对基于移动互联网短视频平台的青睐,短视频弹幕成为一种新兴的互联网互动形式。短视频弹幕是一个在短视频播放过程的实时文字显示系统。用户通过屏幕上的移动评论来表达对当前视频的情绪和看法,也可通过弹幕与其他观看视频的用户进行交流和互动。短视频弹幕可以给观众一种“实时互动”的错觉,虽然不同弹幕的发送时间有所区别,但其只会在视频中特定的一个时间点出现,因此在相同时刻发送的弹幕基本上也具有相同的主题,在参与评论时就会有与其他观众同时评论的错觉。短视频弹幕因其互动性、即时性、娱乐性等特点,受到越来越多用户的喜爱,成为了短视频平台的重要组件之一。

与普通的视频评论相比,短视频弹幕具有更强烈的情感倾向,蕴含丰富的情绪价值^[1]。因此,弹幕的情绪识别与预测极具价值且富有挑战。短视频弹幕情绪识别预测是通过对弹幕文本内容进行分析处理,判断电商用户对短视频的情绪状态。一方面可以帮助视频创作者和平台运营商更好地了解用户对视频内容的反应,从而进行针对性的内容生产和平台运营。另一方面也可以为舆情分析、广告投放等领域提供技术支持。其挑战性在于,弹幕文本内容存在的大量口语化、网络用语以及表情符号等特殊表达方式,给文本的分析带来极大困难。同时,还需要克服视频类型多样性和用户情绪多样性的差异,达到更准确和稳定的预测效果。

目前,短视频弹幕情绪识别预测方法主要分为2类:浅层学习模型解决文本分析问题;通过人工标注获得样本特征,利用机器学习对文本进行分类并提取特征,在训练集的基础上对测试集进行预测。例如,Ellery等学者采用众包结合机器学习的方法,对众包的一小部分语料进行人工标记获取人身攻击数据集,利用这些数据训练了一个机器学习分类器,通过对特征和标记方法进行实验,分类器的阈值在评论子集上产生的结果与人工标记的结果相当^[2]。Jayadev在对评论数据集预处理基础上构建词频矩阵,利用L1正则化逻辑回归和逻辑斯蒂回归对侮辱性评论进行判断,测试发现模型具有很好的侮辱性评论判断效果^[3]。虽然浅层学习模型具有良好的精度,但大多数机器学习方法是基于统计的方法,不仅获取的文本特征准确度低并且比较松散^[4]。

深度学习模型解决分类精度不稳定问题:利用深度学习算法获取文本特征,并使用分类器进行文本分类。例如:Howard等学者提出一种无监督语言模型的微调模型(ULMFIT),与ImageNet模型相似,利用3层BiLSTM捕获上下文语义信息,实现了针对任何自然语言处理任务的迁移学习,在小样本分类上有非常准确的效果^[5]。Alec等学者提出了异构图注意力网络框架(HAN),利用单一任务不可知模型实现强的自然语言理解,并使用无监督预训练显著提高了任务分类的性能^[6]。

2 短视频弹幕处理方法

2.1 短视频弹幕获取方式

以快手平台作为短视频弹幕的来源,快手平台用户群体广泛,影响力大。用户发送的短视频弹幕信息都存储在每个视频的弹幕池中,通过发送网络请求来获取历史播放量最高的50个短视频的弹幕信息。由于弹幕信息都是以XML格式存储的,需要使用Python中的XML模块解析弹幕信息。

2.2 弹幕文本预处理

短视频弹幕文本数据通常包含很多无用信息,通过文本预处理能够清理掉绝大部分。文本一般包括数据清洗,中文分词和去停用词三个步骤。这里展开讨论如下。

(1)采用正则表达式的方法,根据特定规则匹配并删除弹幕文本中的无关信息,例如空格符、换行符、表情符号、特殊字符等,以确保得到的数据仅包含有效文本信息。

(2)采用哈希表的方法对每个弹幕进行哈希处理,并将其键值存入哈希表中,当出现重复弹幕时,只保留第一次出现的弹幕数据,避免重复统计。

(3)随后需要进行中文分词,将弹幕文本数据传入jieba分词库进行分词,有助于后续弹幕的分析。

(4)对于分词后的文本,使用自定义的停用词库,将其中的停用词在分词结果中去掉,例如:“的”、“了”等,减小数据规模。

2.3 弹幕情绪分类法

短视频弹幕的情绪维度可大致划分为正向、中立和负向三大类。其中,正向和负向情绪可通过2个基本维度来描述:愉悦-不愉悦维度和唤起-镇定维度,这2个基本维度构成了情绪空间的四大象限:愉悦-唤起的快乐情绪、愉悦-镇定的满足情绪、不愉悦-镇定的沉默情绪以及不愉悦-唤起的难过情绪。这里愉悦-不愉悦反映的是情绪带来的快乐或

难过程度。唤起-镇定反映了情绪所带来的身体反应和心理状态。中立情绪在研究中被视为短视频弹幕缺乏情感或者反应的状态,但是现实中中立情绪是指一种没有明显情感倾向或者情感表现的状态,表现得比较客观^[6]。情绪空间如图 1 所示,情绪状态随箭头逐渐降低。

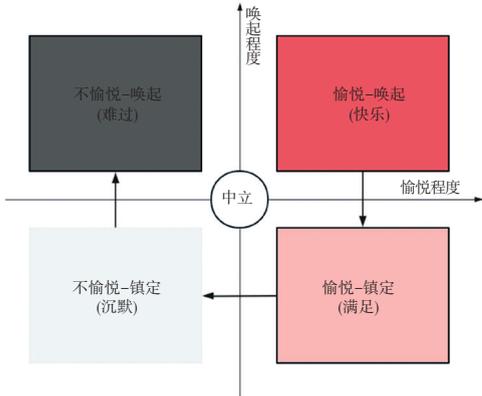


图 1 短视频传播时受众的情绪空间

Fig. 1 The emotional space of the audience during the dissemination of short videos

3 短视频受众情绪预判模型

3.1 深度学习

BiLSTM 和 BiGRU 是机器学习和自然语言处理领域常用的深度学习模型,这 2 种模型在考虑前向和后向上下文信息的同时,具有一定的泛化能力。但是这 2 种模型存在一些不可避免的问题,例如:2 种模型中参数较多,在数据集较小的情况下,模型容易过拟合,对序列中关键信息关注不足等。因此本文提出了一种新的深度学习模型,模型结合 BiLSTM 和 BiGRU,并引入注意力机制更好地捕获关键信息。基于以下考虑:BiLSTM 具有较好的考虑输入序列上下文信息的能力,同时结合 BiGRU 在计算性能上的优势,增加模型建模能力,尤其是处理长序列能力的同时,提高模型训练和推理的效率。基于注意力机制的 Bi(LSTM+GRU)模型的工作流程如图 2 所示。

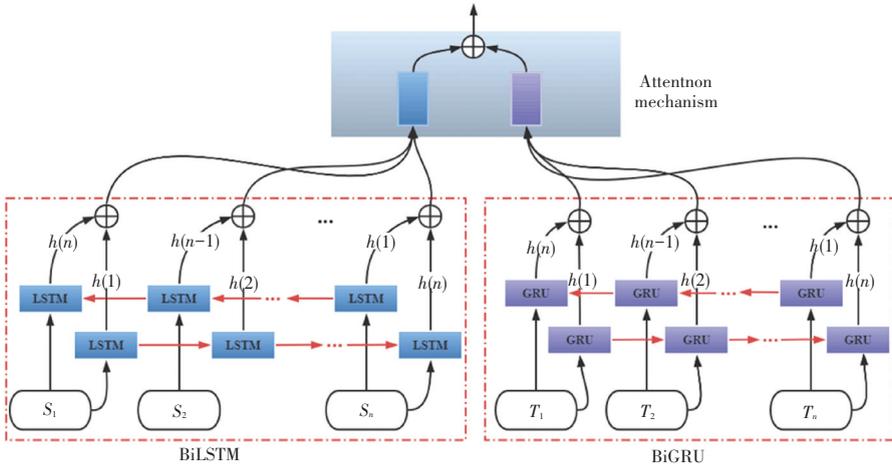


图 2 短视频受众情绪预判深度学习模型

Fig. 2 Deep learning model for predicting the emotions of short video audience

模型首先使用 BiLSTM 和 BiGRU 对输入序列 $S_1 \sim S_n, T_1 \sim T_n$ 进行编码,得到 2 个输出向量 \mathbf{h}_{LSTM} 和 \mathbf{h}_{GRU} 。然后通过注意力机制对每个时间步的注意力权重进行计算,并将加权后的输出向量 \mathbf{Y} 作为模型的最终输出。计算过程如下:

$$\mathbf{H}_{con} = [\mathbf{h}_{LSTM}, \mathbf{h}_{GRU}] \in R^{batch_size \times seq_len \times 2 \times hidden_size} \quad (1)$$

式(1)表示将 BiLSTM 与 BiGRU 的输出进行拼接,得到一个维度为 $(batch_size, seq_len, 2 \times hidden_size)$ 的张量。

研究中对拼接后的张量 \mathbf{H}_{con} 进行线性变换,得

到一个维度为 $(batch_size, seq_len, attention_dim)$ 的张量,计算公式如下:

$$\mathbf{H}_{linear} = \mathbf{T}_a \mathbf{H}_{con} + \mathbf{b}_a \in R^{batch_size \times seq_len \times attention_dim} \quad (2)$$

其中, $attention_dim$ 表示注意力机制中的特征维度; $\mathbf{T}_a \in R^{2 \times hidden_size \times attention_dim}$, $\mathbf{b}_a \in R^{attention_dim}$ 分别表示线性变化的权重和偏置。

研究中也可对变换后的张量进行 Softmax 操作,得到一个维度为 $(batch_size, seq_len, 1)$ 的注意力分数张量,计算公式如下:

$$\alpha_i = \frac{\exp(H_{\text{linear},i})}{\sum_{j=1}^{\text{seq_len}} \exp(H_{\text{linear},j})} \in R^{\text{batch_size} \times \text{seq_len} \times 1} \quad (3)$$

其中, α_i 表示第 i 个位置对应的注意力权重。

研究还将注意力分量与原始 BiGRU 输出按元素相乘,随后对加权 BiGRU 输出张量进行加和,推得的公式分别如下:

$$H_{\text{weight},i} = h_{\text{GRU}} \times \alpha_i \in R^{\text{batch_size} \times \text{seq_len} \times \text{hidden_size}} \quad (4)$$

$$H_{\text{BiGRUsum}} = \sum_{i=1}^{\text{seq_len}} H_{\text{weight},i} \in R^{\text{batch_size} \times \text{hidden_size}} \quad (5)$$

$$H_{\text{BiLSTMsum}} = \sum_{i=1}^{\text{seq_len}} H_{\text{weight},i} \in R^{\text{batch_size} \times \text{hidden_size}} \quad (6)$$

同理得到加权 BiLSTM 输出张量和,将 2 个输出张量进行拼接,得到新的输出矩阵 H_{sum} 。最后得到对于第 i 个词的重要性:

$$e_i = \alpha_i \tanh(W_{\text{sum}} H_{\text{sum}} + b) \quad (7)$$

其中, W_{sum} 表示一个线性变换的权重矩阵; b 表示偏置向量; \tanh 表示双曲正切函数; α_i 表示注意力权重。最后计算加权后的输出向量为:

$$Y = \sum_{i=1}^n W_i h H_{\text{con},i} \quad (8)$$

其中, n 表示序列长度。

3.2 浅层学习模型

(1)Fast Text 预训练。短视频弹幕文本经过预处理后,输入到 Fast Text 层进行文本数据的标签化处理,将每个文本对应的类别标签映射为整数表示,随后将文本中的词汇表示为文本向量。Fast Text 预训练层使用字符级别的 N-Gram 特征,可以很好地处理罕见词汇,在快速处理大规模文本数据的同时,保证了文本分类的准确率。

(2)决策树分类。经过深度学习模型对文本特征的提取后,将提取的特征输入到浅度学习模型中进行分类,可以利用深度学习模型强大的特征提取能力和浅度学习模型的高效分类能力。这里采用决策树模型用于模型最后的分类,由于决策树模型可以提供一定的正则化效果,从而减轻深度学习模型的过拟合问题。

3.3 集成模型

本文基于上述的深度学习和浅度学习的分析,用图 3 描述了 2 种异构分类器的集成模型。由图 3 可知,短视频弹幕文本数据经预处理后进入 Fast Text 层进行文本数据的预训练,生成更易被深度学习模型学习的文本向量。文本向量输入到前向 GRU、LSTM 和后向 GRU、LSTM 层中,对输入文本特

征进行训练。经过 Bi(GRU+LSTM) 训练后得出的特征向量作为 Attention 层的输入,对特征向量进行加权后,将 Attention 层获得的语义向量输入到决策树层,最终实现短视频弹幕情绪识别预测。

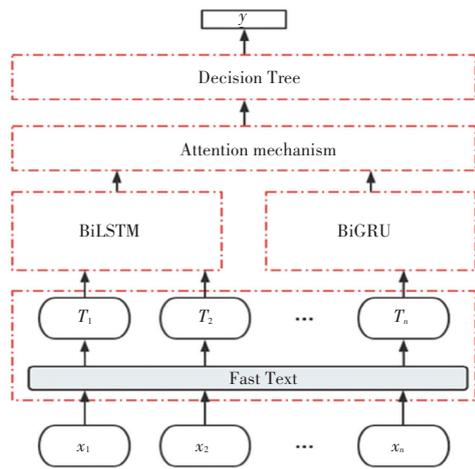


图 3 情绪预测集成模型

Fig. 3 Integrated model for emotion prediction

4 受众情绪预判实证分析

将数据集按照 2 : 8 的比例划分为测试集和训练集。设置 BiLSTM、BiGRU、BiGRU-Att、BiLSTM-Att 共 4 个模型进行对比实验,模型均使用 Fast Text 预训练模型进行文本向量化,迭代训练 30 次。其中,Fast Text 词向量层设置 dim 为 120, $maxn$ 指定 N-Gram 最大长度为 5,使用 Sigmoid 激活函数将词向量压缩到 0~1 范围内。BiLSTM(BiGRU) 使用 2 层 LSTM(GRU),隐藏层尺寸为 128, dropout 为 0.1,使用 2 个全连接层。第一个全连接层神经元个数为 100,第二个全连接层使用激活函数 Sigmoid。

各类深度学习模型性能对比见表 1。从表 1 数据可以看出,当模型被引入注意力机制后,其性能相比不引入注意力机制来说也会明显提高。在迭代 30 次后,Bi(LSTM+GRU) 模型的性能明显优于其它模型。在最后将浅度学习模型决策树作为结果输出的文本标签分类时,模型的性能往往会向优于原模型方向发展,但效果不明显;当原模型的过拟合问题比较严重时,这个模型的性能可能表现得更好。通过观察发现,所有模型在训练集和测试集的准确率和损失函数都明显偏优,但是模型的精准率却不是理想,分析主要原因在于短视频弹幕部分为文字加表情结构,一个弹幕往往因为一个表情从而使得整个弹幕的情绪完全转变^[7]。在本文收集的短视频弹幕数据中摒弃了特殊字符,所以会导致情绪识别的准确率不是很高。

表 1 各类深度学习模型性能对比

Table 1 Performance comparison of various deep learning models

模型	训练集损失函数值	训练集准确率/%	测试集损失值	测试集准确率/%	精准率/%
BiLSTM	0.048 6	96.93	0.047 9	96.81	82.9
BiGRU	0.045 5	97.79	0.044 3	97.64	83.6
BiGRU-Att	0.020 1	99.06	0.021 3	99.01	85.2
BiLSTM-Att	0.029 3	98.15	0.024 5	98.67	84.7
Bi(LSTM+GRU)-Att	0.010 9	99.69	0.011 4	99.61	87.6
Bi(LSTM+GRU)-Att-DT	0.010 7	99.70	0.010 9	99.69	87.7

使用模型对预先设计的文本进行情绪识别的预测,预测结果见表 2。从表 2 可以看到,模型能够对短视频弹幕情绪进行准确识别预测。

表 2 模型验证

Table 2 Model verification

短视频弹幕文本	模型情绪识别预测结果
这次专家们又会如何发表看法呢	愉悦-镇定
华莱士最好用	愉悦-镇定
你怎么能这么说?	不愉悦-镇定
红薯,红薯红薯!重要的事情说三遍!	愉悦-唤起
你真敢收钱啊!	不愉悦-唤起
不喜欢北方的天气	中性
您没事吧?	不愉悦-镇定
我不理解,你到底要什么?	不愉悦-唤起

5 结束语

本文提出了一种应用于短视频弹幕情绪识别预测的异构分类器集成方法。该方法通过 Fast Text 获取文本向量,结合双向 LSTM 网络和双向 GRU 网络,引入注意力机制,最后通过浅度学习获得最终的短视频弹幕情绪分类。在情绪分类的类别上,提出一种情绪空间象限,用 4 个不同的象限描述情绪的倾向和反应程度。相对于以往的模型,该模型使用预训练模型进行词向量获取,使得词向量表达的语义特征更强;双向 LSTM 和双向 GRU 网络使得信息能够关联上下文间的语义特征,弥补相距较远的词

依赖错误;注意力机制又能把不重要的语义信息进行隔离,提高情绪弹幕预测和分类的精度。实验表明,在双向 LSTM 和双向 GRU 网络中结合注意力机制的模型表现最优,其性能相比其他模型有明显提升。最后通过输入待预测弹幕验证实验结果,表明模型能够较好地完成短视频弹幕的识别预测任务。

参考文献

- [1] 王志豪. 基于 FCM 的模糊聚类算法研究[D]. 厦门:厦门大学, 2020.
- [2] 张鑫. 基于深度融合先验的模糊聚类算法研究[D]. 南京:南京邮电大学, 2023.
- [3] 陆海青. 基于模糊聚类的图像分割算法的研究与应用[D]. 无锡:江南大学, 2018.
- [5] 王彦, 马丹, 张国辉. 分层次教学的模糊聚类分组模型[J]. 大庆师范学院学报, 2012, 32(3): 31-33.
- [6] 江涛, 黄昌昊, 孙斌. 基于文本挖掘的弹幕情绪分析研究[J]. 智能计算机与应用, 2022, 12(8): 60-64, 69.
- [7] 朱峰. 短视频的传播特点以及传播效果提升策略研究[J]. 电视技术, 2021(10): 24-27.
- [8] 付凡格. 融媒时代都市报新闻短视频传播矩阵的构建及运行研究[D]. 武汉:武汉体育学院, 2022.
- [9] 彭静. 网络直播背景下服装电商发展现状及营销策略分析[J]. 鞋类工艺与设计, 2021(18): 29-31.
- [10] 高爽. 中部六省高质量发展水平测度及比较差异分析—基于投影寻踪模型的评价模型[J]. 太原城市职业技术学院学报, 2023(2): 15-18.
- [11] 白梅, 陆啸尘. 基于用户情绪状态的网络舆情传播模型研究[J]. 智能计算机与应用, 2022, 12(3): 147-151, 157.
- [12] 顾亦然, 王雨, 杨海根. 基于用户行为序列的短视频用户多行为点击预测模型[J]. 电子信息学报, 2023, 45(2): 672-679.