

基于弹幕评论的在线知识社区用户关注度与情感度联合分析*

李佳轩 储节旺

(安徽大学管理学院, 合肥 230601)

摘要: 试图分析知识社区用户弹幕评论中的关注度与情感度之间的关系, 并尝试构建一种关联二者的坐标系, 以此为知识视频的分类与推广提供新方法。通过获取一定数量的弹幕评论, 利用Sentence-LDA模型与XLNet预训练模型对弹幕评论文本进行分析。通过抽取视频弹幕评论的主题和关键词, 对其进行情感估计。根据用户对不同弹幕主题的关注度和情感度, 建立用户关注度-情感度坐标系。结果表明, 随着视频的播放, 在线知识社区用户对视频中不同主题的关注度和情感度发生明显变化。用户对视频的不同主题有不同的情感评价, 即使是对同一主题, 用户也会随着视频的播放而改变情感态度。依据情感与主题的演变, 可以构建新的知识视频内容标签。

关键词: 在线知识社区; 弹幕评论; 联合分析; 知识推广; 视频分类; 情感度分析

中图分类号: G250.7 **DOI:** 10.3772/j.issn.1673-2286.2023.08.008

引文格式: 李佳轩, 储节旺. 基于弹幕评论的在线知识社区用户关注度与情感度联合分析[J]. 数字图书馆论坛, 2023 (8): 68-76.

在线知识社区也被称为在线创新社区或开放式创新社区, 是社区用户分享知识和想法的平台^[1-2]。在线知识社区的建立主要基于社区用户的共同兴趣和爱好^[3-4]。一个在线知识社区通常有诸多版块, 每个版块专注于不同的知识主题, 如自然、战争、人文等^[5]。社区用户可以通过在线知识社区分享他们的知识, 同时, 用户可以对他人分享的知识进行评价^[6]。目前, 美国的Quora和Stack Overflow以及中国的知乎等在线知识社区已经成为重要的知识分享平台。这类社区的成功源于其具备提供高质量知识以满足用户认知需求的能力^[7]。与传统的社会社区不同, 在线知识社区专注于创造和传播知识, 而不是个人情感、新闻或评论^[8]。

随着在线知识社区服务功能的完善, 社区用户分享知识的方式不再局限于文字和图片。弹幕视频作为一种交流知识和表达思想的方式已经逐渐成为一种新的

趋势, 可以让用户在观看视频的同时进行即时交流^[9]。在弹幕视频中, 用户可以让他们的评论出现在视频的某一帧中, 其他用户在观看到该帧时可以看到此评论。当一个视频有很多评论时, 就会产生无数评论飞过屏幕的效果, 这种现象被称为“弹幕”。弹幕视频中的用户评论并不会显示在评论区, 而是被锁定于视频中的某个时间段, 供以后的观众观看, 这创造了一种共同观看视频的感觉^[10]。弹幕评论功能起源于日本的在线视频平台Niconico, 现在该功能已经成为中国在线知识社区的基本特征之一^[11]。在在线知识社区中, 观看知识视频的用户经常通过发送弹幕评论来交流知识, 用户还可以通过选择字体大小、颜色和位置来实现个性化评论^[12]。

作为在线知识社区中一种新的知识交流形式, 弹幕评论具有实时性强、存在感强的特点^[13]。用户的评论出现在视频某一帧的画面, 可以更准确地反映用户

收稿日期: 2023-06-06

*本研究得到安徽省哲学社会科学规划重大项目“安徽打造具有重要影响力的科技创新策源地研究”(编号: AHSKZD2021D02)资助。

观看视频时的态度。弹幕评论功能允许众多知识社区用户同时发布内容, 用户也可以利用弹幕评论功能进行实时交流, 这在一定程度上促进了用户之间的知识传播和交流。以往研究多分析弹幕评论的主题内容或情感特征, 极少有学者探讨两者之间的关系。本研究试图以联合分析的方法, 将弹幕评论的主题与情感进行关联, 构建一种基于弹幕评论关注度与情感度的坐标系, 并以坐标的形式提取知识视频的核心内容与用户情感演化趋势, 从而为知识视频的推广与分类提供新思路。

1 在线知识社区弹幕评论的相关研究

1.1 不同领域对弹幕评论的探讨

近年来, 与弹幕评论相关的研究数量逐渐增多, 不同学科的学者提供了多种研究视角。在传播学领域, 弹幕评论通常被认为是一种新的传播模式。一些学者认为, 这种新的传播模式可以有效地促进受众之间的沟通^[14]。学者还认为, 这种传播模式可以为平台带来新的收入^[15]。在信息科学领域, 用户发送弹幕评论的原因一直是学者关注的主要方向。一些学者尝试使用S-O-R (Stimulus-Organism-Response) 理论来研究用户发送弹幕评论的动机。根据SOR理论, 学者发现用户在发送弹幕评论时会感受到明显的社会存在感^[16]。除此之外, 对自我表达的渴望也是用户发送弹幕评论的主要原因^[17]。随着越来越多的平台使用弹幕评论, 发送弹幕评论已经成为一种青年文化, 用户会出于一些更新奇的原因而发送弹幕评论。在知识社区中, 用户发送弹幕评论的主要原因是他们的知识需求^[18]。在社会学领域, 学者更多关注弹幕功能背后的文化现象。弹幕文化蕴含着不同时期人群的集体特色^[19], 因此也有学者基于弹幕评论文本挖掘社会中的群体现象与极化现象^[20]。

研究表明, 弹幕评论远比普通的评论方法更能体现用户当下的情感和态度^[21]。用户经常会选择发布弹幕评论而不是在评论区中表达他们对视频的看法。此外, 弹幕评论具有明显的时效性与瞬时性, 因此在文本研究中, 弹幕评论有助于学者了解用户即时的情感波动和行为表征。

1.2 弹幕评论中用户关注度的相关研究

在线知识社区中, 用户通过分享、评论来表达对视

频的态度, 而用户分享与评论等行为的频次也能够体现用户群体对视频的整体关注度, 但是该类指标缺乏时效性, 同时难以突出用户的情感倾向。对此, 有学者尝试从弹幕的视角出发, 利用弹幕评论的瞬时性特点挖掘用户独特的信息行为^[22]。目前, 关于弹幕评论中用户关注度的研究主要分析用户表达关注的动因、量化用户的关注行为。已有研究认为用户在弹幕评论中表达关注的动因主要为社区文化与视频主题以及用户自身的兴趣爱好^[23]。除此之外, 弹幕评论的发布效率高、情绪表达简易化等特点也是用户愿意在弹幕评论中表达关注的原因^[24]。量化用户的关注行为是指对弹幕中的短文本进行数据分析, 包括主题分析、时序分析等。王文韬等^[25]通过提取健康视频弹幕评论, 剖析用户对健康视频不同主题的情感。黄立赫等^[26]通过分析弹幕评论中的关注度预测网络舆情。在这类研究中, 学者普遍使用主题分析的方法构建弹幕视频中的用户关注度指标。然而, 在线知识社区中的用户关注度既包括由弹幕评论的主题传达出的用户对视频的整体关注度, 也包括不同时刻用户对视频不同主题的关注度。前者偏向用户的整体评价, 而后者则考虑时序变化、情感波动以及视频播放过程中用户对视频不同主题的态度变化。

通过弹幕评论挖掘用户对视频主题的关注度的研究并不少见, 有学者发现, 弹幕评论不仅可以体现用户对视频内容的关注度, 也可以用来对用户进行分类^[27]。然而, 仅依靠关注度研究无法有效、准确地分类用户、挖掘知识视频内容, 因此情感度分析成为了学者关注的方向。

1.3 弹幕评论中用户情感度的相关研究

表达情感是用户发布弹幕评论的重要原因。对于知识视频的用户来说, 弹幕评论是他们评价知识视频的主要方式。对于知识视频的创作者而言, 则可以根据用户弹幕评论找到视频的亮点与痛点。

在以往的研究中, 学者通常根据不同时间段的弹幕评论数量分析用户的情绪波动^[28], 但是, 弹幕评论数量的变化并不能准确反映用户情绪波动, 因此弹幕评论的文本内容是当前用户情感度分析的主要依据。目前学者通常使用机器学习、情感词典等方法来分析弹幕评论的内容。例如, Yamamoto等^[29]将弹幕评论中的形容词和语气词作为特征, 通过使用支持向量机分类器实现对视频的情感度分类。Hong等^[30]使用情感词典来

计算用户的正面和负面情感度,对用户进行分类。这些研究可以有效地帮助知识社区的管理者对视频和用户进行分类,这也是目前大多数情感度研究的主要目的。在我国,视频弹幕评论情感度研究主要用于防止极化现象与社会舆情。

在本研究中,弹幕评论的情感度主要被区分为用户在弹幕评论中的整体情感度与用户对不同主题的情感度,并以量化评分的方式进行计算。值得关注的是,弹幕评论最重要的特点是时效性与瞬时性,这使得用户评论的关注度与情感度会随着时间的推移而变化。换句话说,弹幕评论的关注度与情感度并不是固定的。因此,本研究基于前人的研究成果,将弹幕评论中用户对不同主题的关注度与情感度进行联结并计算用户关注度与情感度的具体值,最终建立包括用户关注度和情感度的坐标系,以立体的形式展现用户对某一视频的整体态度。同时,引入时间变量以揭示弹幕评论中用户关注度、情感度随时间的变化。

总的来说,本研究将实现以下目标:①实现对弹幕评论中用户关注度与情感度的关联挖掘;②分析弹幕评论中用户关注度与情感度的动态演化趋势;③为在线知识社区中知识视频的标签标记与分类推广提出新思路。

2 研究方法

2.1 弹幕评论中关注度的分析方法

随着自然语言处理技术的普及,用户评论文本研究领域已有非常丰富的成果^[31-32]。弹幕评论作为一种新的短文本评论模式,受到了学者的关注。目前涉及短文本中用户关注度的研究大多采用Latent Dirichlet Allocation (LDA)模型,这类研究试图通过主题建模的方式挖掘用户对不同视频或新闻的关注方向,以主题表征用户的关注方向,例如使用LDA模型来研究学生的聊天记录、求职者的介绍信等短文^[33]。自开发以来,LDA模型经过多次优化,一些新的LDA模型如BERTopic模型等被提出。由于弹幕评论具有明显的短文本特征,选择Sentence-LDA模型来分析弹幕评论中的用户关注度。

Sentence-LDA模型考虑了文档与主题、主题与词之间的多指标分布。一个文档可能有多个主题,而主题包含多个关键词。学者常使用困惑度确定主题的数量。困惑度是指每个句子出现概率的几何平均值的倒数,也就是说,困惑度越低,可选择的主题数量越少,模型

也就越准确。在具体研究过程中,利用弹幕评论的主题表征用户关注度,以权重为关注度的具体数值。同时根据视频时长对弹幕评论文本进行切分,确定不同时间段用户关注度的演化情况,并进行情感度分析。

2.2 弹幕评论中情感度的分析方法

情感度分析是基于文本判断用户情感的过程,包括从文本中提取情感信息,进行情感评价。情感度分析的基本任务是情感识别和情感计算。目前主要有3种分析情感度的方法:基于机器学习的方法、基于词典的方法和混合方法。对于弹幕评论中的情感度分析,学者开展了一些研究,如:孙晓宁等^[34]使用情感词典计算用户的正面和负面情感度,并根据情感度对用户进行分类;Hong等^[30]使用情感词典来识别弹幕评论中的用户情感度,并根据情感度的波动来判断视频的亮点。也有许多学者利用机器学习方法分析弹幕评论中的用户情感度^[35]。在已有研究中,学者通常以弹幕评论中的整体情感度判别情绪倾向,忽视了用户关注度与情感度的关系,同时用户的情感度会随着视频的播放而产生波动。

目前,较为优秀的情感度分析模型如BERT、XLNet等极少用于弹幕视频中的情感度分析。相比BERT模型,XLNet采用了Permutation语言模型以及双流自注意力机制,结合了Transformer-XL的相对位置编码,在文本翻译、情感识别方面的效果优于BERT模型^[36]。但是,XLNet缺点是只能利用上文或者下文的信息,不能同时利用上文和下文的信息。不过对于弹幕评论这样的瞬时短文本,该缺点并不会显著影响研究结果。因此,选择使用XLNet预训练模型。在具体实验过程中,分别计算用户对弹幕视频的整体情感度以及对不同主题的情感度,并以量化的形式展示。

2.3 关注度-情感度坐标系构建

根据确定的用户关注度与情感度,构建弹幕视频中的用户关注度-情感度坐标系。该坐标系以横轴表示用户对弹幕视频中主题的关注度,以纵轴表示用户对该主题的情感度。在此坐标系中,分别设定高关注度(情感度)、中关注度(情感度)、低关注度(情感度)3个维度,以此区分用户的关注度与情感度。通过该坐标系既能够体现用户对弹幕视频的整体关注度与情感度,也能够体现用户对视频不同主题的关注度与情感

度,从而细粒度挖掘用户的潜在行为特征。同时该坐标系也有助于在线知识社区对弹幕视频进行分类与推广,从而为用户提供更具针对性的知识视频。

3 实验过程与结果

3.1 数据获取与预处理

为了保证研究的可靠性与文本数据的充足性,选择bilibili平台作为研究对象。bilibili是一个综合在线知识社区,其中存在诸多版块,如动漫、教育、科学、天文地理、人文历史等,其特色就是发布弹幕视频^[37]。在bilibili平台中,历史版块拥有较多的关注者。使用Python工具获得于2020年1月1日—2022年4月1日发布的时长约为15分钟的有关清朝历史的视频,得到共计65 201条弹幕评论。实际分析中发现,由于存在奖励机制,用户会发送一些无意义的词语,如“快点”“冲”“123”等。此外,获得的数据中也有很多重复和缺失的数据。因此,建立一个全面的用户词库和停用词表,并对获得的数据进行预处理,具体操作包括词语分离、去除重复和无效的信息和表情符号,最终得到31 500条有效的弹幕评论。

3.2 弹幕评论的关注度分析

根据经过处理的弹幕文本,以主题模型分析弹幕

评论中用户关注的主题,从而确定弹幕评论中的用户关注度,即以主题表征用户的关注度。实际分析过程中发现,随着视频的播放,弹幕评论的主题也在发生变化,根据这一特征计算不同弹幕评论主题的权重和弹幕评论主题的变化轨迹。弹幕评论主题的权重作为用户关注度的指标,用于后续建立关注度-情感度坐标系。

根据上述逻辑,使用Sentence-LDA模型来分析弹幕评论文本中的用户关注度。将模型的主题数量设置为1 000个,超参数为0.000 1,最小误差率为0.01,迭代次数为40次,每个主题下的关键词数量为1 800个^[38]。随着主题数量的增加,困惑度先降后升(见图1)。当主题数量为7个左右时,困惑度最低,因此最佳主题数量为7个。提取出的权重较高且与视频相关度较高的7个主题如表1所示。

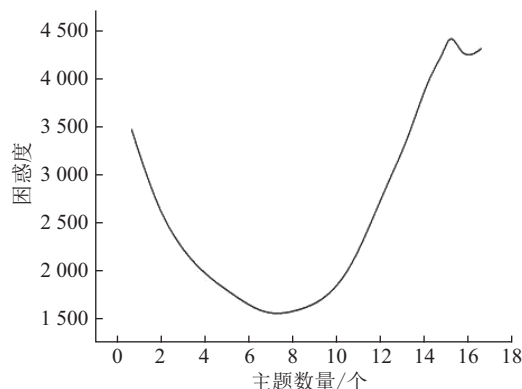


图1 不同主题数量下的困惑度

表1 弹幕评论的主题与关键词

序号	主题	关键词
0	清朝经济	工业、劳动力、高附加值、商业、出口、货币、资本
1	清朝战争	海战、武器、八国联军、第一次中日战争、鸦片战争、枪炮
2	人民反抗斗争	民族、历史、抵抗、爱国、太平天国、义和团运动、剥削
3	清朝条约	国家、北京条约、外债、白银、北京、南京条约、资源
4	清朝外交	武士、幕府时代、日本、王子、天皇、马关条约、东洋
5	清朝统治者	康熙、雍正、慈禧、朝代、溥仪、顺治、同治
6	清朝官员	李国忠、晚清、军阀、左志东、曾国藩、林志鑫

研究选择的视频主要有关清朝历史,所以大部分的弹幕评论内容与清朝高度相关。如表1所示,知识社区用户发布的弹幕评论主要涉及清朝的经济状况、清政府的战争、人民的反抗活动、清政府的外交条约、清朝的统治者、清朝的官员,即用户对此类主题具备较高的关注度。例如,在“人民反抗斗争”主题中,清朝人民

的反抗活动是用户关注的焦点。为了探索用户发布的弹幕评论的主题以及主题的变化趋势,具体分析用户弹幕评论的关键词。首先,根据已提取的主题,基于视频播放的时间,每隔1分钟对与主题相关的关键词进行统计。其次,根据表1中的关键词对统计结果进行修正。最后,选择一些关键词进行展示,见图2。

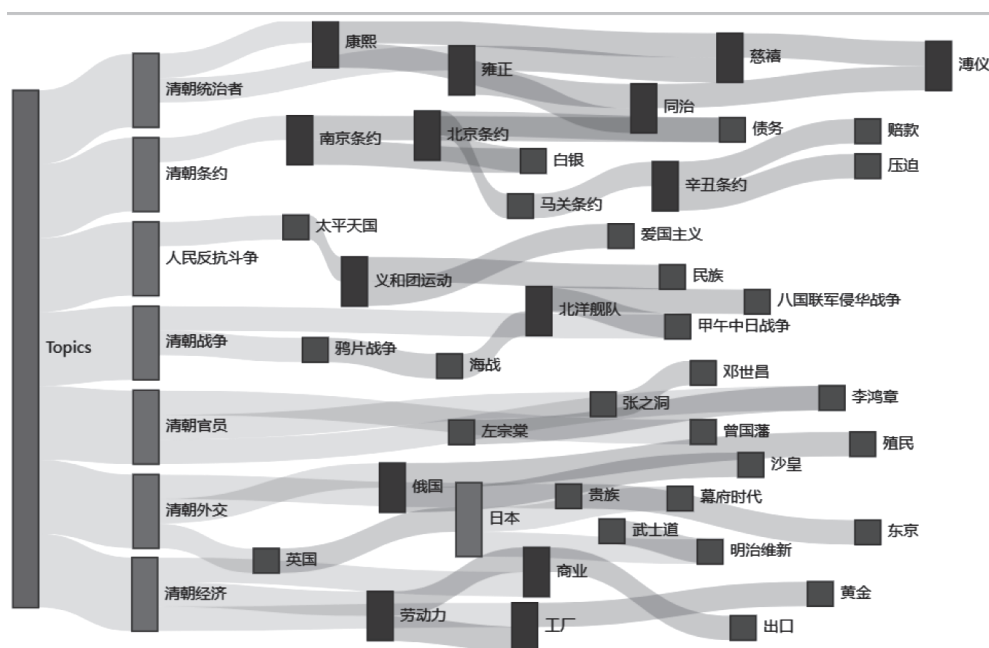


图2 弹幕评论的主题变化

以“清朝条约”这一主题为例，在视频播放至第5分钟时，许多用户对《南京条约》发表了自己的看法。当视频播放至第11分钟时，许多用户发表了与《马关条约》有关的弹幕评论。“南京条约”“马关条约”等关键词共同构成了“清朝条约”这一核心主题。

3.3 弹幕评论的情感度分析

3.3.1 弹幕评论的整体情感评价

主要通过计算弹幕评论的情感分值来分析知识社区用户对视频的情感度。采用的训练数据包括谷歌公司提供的训练数据、大连理工大学信息检索研究室开发的情感词表以及百度为弹幕评论开发的多维情感词典^[39]。在研究过程中，将所需预测的弹幕评论文本以句为单位循环输入XLNet模型。模型采用二分类方法对情感进行评价：0表示最悲观的评论，1表示最乐观的评论。研究发现，用户的弹幕评论中存在一些流行的社区术语，这些术语也具有明显的情感倾向，如“2333”“高燃”“高能”等，将这些术语替换为同类型中文词语，以此提高情感预测的准确度。

弹幕评论的情感分值具体结果如图3所示。图3显示，情感分值为0或1的弹幕评论数量最多，这说明知识社区用户对清朝历史的看法存在极化现象。同时，情感分值为0.5的弹幕评论数量也比较多，这说明很多知识

社区用户对清朝历史的评价并没有太多感情倾向。用户对视频的平均情感分值为0.647，说明用户对视频内容的整体评价偏向正面。

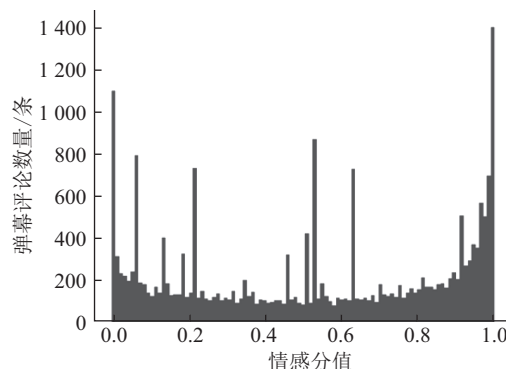


图3 弹幕评论的情感分值

3.3.2 不同弹幕评论主题的情感度分析

在以往的研究中，学者倾向于根据知识社区用户对视频的情感评分来为视频添加情感标签^[40]，但这些研究并没有考虑用户对不同的主题有不同的情感度。因此，试图从弹幕评论中不同的主题出发，分析用户对同一主题的情感度波动和对不同主题的情感度变化。重点根据图2所示的关键词，计算用户对同一主题下不同关键词的情感度，以此判别用户多样的情感度变化。以主题“清朝统治者”为例，分别在弹幕评论中提取含有“慈禧”以及“溥仪”的评论，并形成两个新的文本，计

算文本情感分值。

根据此逻辑, 计算了用户在不同时间段对同一主题下不同关键词的情感分值, 根据此分值绘制情感演化曲线, 如图4所示。图4中, 横轴为时间段, 纵轴为弹幕评论的情感分值(0.50分为中性评价, 高于0.50分为乐观评价, 低于0.50分为悲观评价)。例如, 在“清朝统治者”主题中, 用户在视频的前7分钟对康

熙和雍正等的评价保持中立, 但在视频的第11分钟左右, 情感分值变成了0.15, 这意味着用户对慈禧的评价极差。

计算用户对不同主题的平均情感分值作为用户情感度, 如表2所示。从表2可以看出, “清朝条约”和“清朝外交”主题的用户情感度较低, 而“清朝官员”和“人民反抗斗争”主题的用户情感度较高。

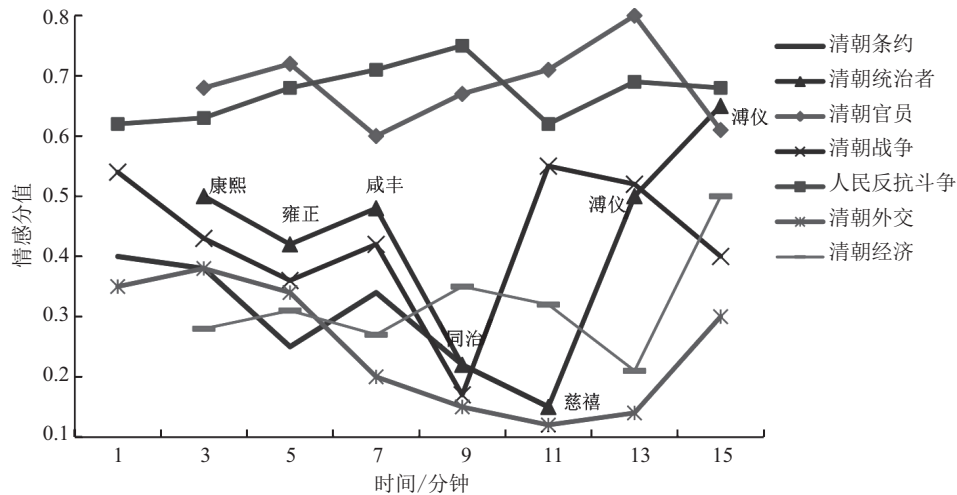


图4 弹幕评论主题关键词的情感分值波动

表2 弹幕评论主题的用户情感度与权重

序号	主题	情感度	权重
0	清朝经济	0.25	0.14
1	清朝条约	0.27	0.17
2	清朝统治者	0.41	0.19
3	清朝官员	0.69	0.09
4	清朝战争	0.44	0.15
5	人民反抗斗争	0.67	0.06
6	清朝外交	0.23	0.18

3.4 在线知识社区用户关注度-情感度坐标系构建

首先, 计算出不同弹幕评论主题的权重, 并将其视为用户对不同主题的关注度。研究得到了7个弹幕评论主题和与主题高度相关的关键词。对于这些关键词, Sentence-LDA模型提供了TF-IDF算法来计算关键词的权重。得到了7个主题的关键词权重, 并选择了不同主题中最大的关键词权重作为弹幕评论主题的权重, 不同弹幕评论主题的权重如表2所示。其次, 分析了用

户对不同的弹幕评论主题的情感分值, 情感分值体现了用户对不同主题的情感度。最后, 根据用户的关注度和情感度来构建坐标系, 以体现用户对某一知识视频的整体观感和态度, 如图5所示。

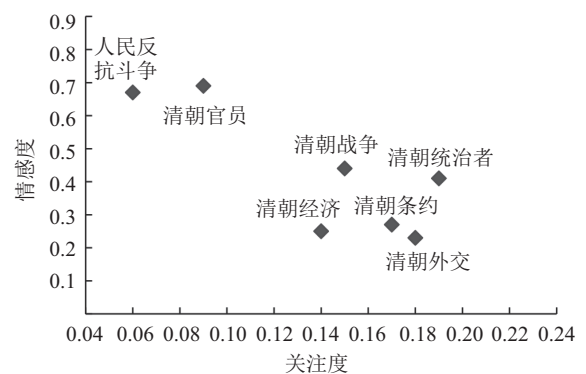


图5 在线知识社区用户关注度-情感度坐标系

根据实验数据与实验结果, 以三等分点为基础, 将用户的情感度分为高情感度(情感分值 ≥ 0.6)、中情感度($0.4 < \text{情感分值} < 0.6$)、低情感度(情感分值 ≤ 0.4)。由于主题权重的差异较大, 根据所有主题的权重均值, 暂且将“清朝官员”“人民反抗斗争”列为高情感度主

题,“清朝统治者”“清朝战争”列为中情感度主题,“清朝经济”“清朝条约”“清朝外交”列为低关注度主题。之后以同样的逻辑划分关注度,例如:“清朝条约”的关注度为0.17,情感度为0.27,属于低情感度、高关注度主题;“人民反抗斗争”的情感度为0.67,关注度为0.06,属于低关注度、高情感度主题。通过该坐标系可立体化描述与评价知识视频,相对于以往简单的视频标签与描述,该坐标系具有以下优点。①不仅能够描绘用户对知识视频内含主题的关注度,还能体现用户对该主题的情感度。这既有助于其他用户提前获取知识视频的相关内容,也有助于为视频作者提供知识传播效果的反馈。②可以作为分类知识视频的新方式,结合情感度与关注度两个维度,为视频提供全新的标签,便于知识视频的推广。

4 结论

弹幕视频作为在线知识社区中较为流行的知识传播方式,允许用户的评论出现在视频的某一帧上。在以往的研究中,虽然有学者将弹幕评论作为分析网络社区用户行为特征的依据,但是很少有学者关注弹幕评论中随着时间不断变化的用户关注度和情感度,也很少有学者将不同主题和不同情感的弹幕评论联系起来。王文韬等^[25]、王晓等^[41]曾尝试将用户的关注主题与情感进行关联,但没有考虑同一主题下存在多个细分主题,同时用户对这些细分主题存在着不同的关注度与情感度。例如,在“清朝统治者”主题下,用户对溥仪有一定的好感,但对慈禧有明显的不满。除此之外,在研究模型的选择上,本研究应用了目前较为流行的预训练模型,而非基于词典的情感识别方法。此外,在与相关研究的对比中发现,相较于在线健康社区与问答社区,在线知识社区的用户呈年轻化趋势,情感表达更为强烈,且使用独特的流行语。因此,辨析社区用语成为在线知识社区用户关注度与情感度研究的关键。

本研究提出关注度-情感度坐标系,其可以成为知识视频的一种标签形式。相对于以往的文本评论或标签标记,立体化的坐标标签更能体现用户对某知识视频内容的关注方向与情感判断。根据关注度指标,知识视频制作者可以了解用户对知识视频中不同主题的关注程度,获得用户的全面反馈,从而找到更新视频的方向;对于在线知识社区而言,关注度指标为知识视频评价提供了一种新的方法。情感度可以显示用户对不同

主题的情感倾向,而在线知识社区可以利用它来为用户推荐相关的视频。

本研究也有一些局限性。不同的在线知识社区有独特的弹幕评论语言,所以研究采用的训练文本不够全面。在关注度方面,根据视频的时间来提取弹幕评论的关键词,在描述主题演变的准确性方面存在一些问题。用关键词的权重来体现用户对弹幕评论主题的关注度,存在维度单一、缺乏准确性的问题。

在未来,将改进方法,通过完善情感词典或使用混合分析的方法进行情感判断。对于用户关注度的变化,将同时考虑弹幕评论的出现时间、视频时长、评论区内容等维度,或使用动态主题演化模型或BERTopic模型进行分析。希望提出的坐标系能够推广到其他领域,如在线健康社区、问答社区等。

参考文献

- [1] WANG N, TIBERIUS V, CHEN X X, et al. Idea selection and adoption by users: a process model in an online innovation community[J]. *Technology Analysis & Strategic Management*, 2021, 33 (9): 1036-1051.
- [2] ZHAO L, DETLOR B, CONNELLY C E. Sharing knowledge in social Q&A sites: the unintended consequences of extrinsic motivation[J]. *Journal of Management Information Systems*, 2016, 33 (1): 70-100.
- [3] CHEN W, WEI X H, ZHU K X. Engaging voluntary contributions in online communities: a hidden Markov model[J]. *MIS Quarterly*, 2018, 42 (1): 83-100.
- [4] CHEN H, NUNES J, RAGSDALL G, et al. Extrinsic and intrinsic motivation for experience grounded tacit knowledge sharing in Chinese software organisations[J]. *Journal of Knowledge Management*, 2018, 22 (2): 478-498.
- [5] 储节旺, 李佳轩. 知识生态系统中知识演化及智慧创生研究: 以知乎社区为例[J]. *情报理论与实践*, 2022, 45 (9): 51-58.
- [6] BIEBER M, ENGELBART D, FURUTA R, et al. Toward virtual community knowledge evolution[J]. *Journal of Management Information Systems*, 2002, 18 (4): 11-35.
- [7] CHOU C H, WANG Y S, TANG T I. Exploring the determinants of knowledge adoption in virtual communities: a social influence perspective[J]. *International Journal of Information Management*, 2015, 35 (3): 364-376.
- [8] JIN J H, LI Y J, ZHONG X J, et al. Why users contribute

- knowledge to online communities: an empirical study of an online social Q&A community[J]. Information & Management, 2015, 52 (7) : 840-849.
- [9] BENEVENUTO F, RODRIGUES T, ALMEIDA V, et al. Video interactions in online video social networks[J]. ACM Transactions on Multimedia Computing, Communications, and Applications, 2009, 5 (4) : 1-25.
- [10] LI J Y. The interface affect of a contact zone: danmaku on video-streaming platforms[J]. Asiascape: Digital Asia, 2017, 4 (3) : 233-256.
- [11] HE Y F, TANG T Y. Recommending highlights in anime movies: mining the real-time user comments “Dan-MaKu” [C]//2017 Intelligent Systems Conference (IntelliSys), September 7-8, 2017, London, UK. New York: IEEE Press, 2018: 319-322.
- [12] 杨阳, 余维杰. 融合弹幕内容特征与行为特征的用户画像研究: 以B站教学类视频为例[J]. 情报科学, 2022, 40 (12) : 161-169.
- [13] 魏来, 王伟洁. 基于互动仪式链理论的视频网站弹幕信息情感分析: 以Bilibili健康科普类视频为例[J]. 情报理论与实践, 2022, 45 (9) : 119-126.
- [14] MA Z H, GE J. Analysis of Japanese animation’s overlaid comment (danmu) : a perspective of parasocial interaction[J]. Chinese Journal of Journalism and Communication, 2014, 8: 116-130.
- [15] LI Z, LI R, JIN G H. Sentiment analysis of danmaku videos based on Naïve Bayes and sentiment dictionary[J]. IEEE Access, 2020, 8: 75073-75084.
- [16] FANG J M, CHEN L, WEN C, et al. Co-viewing experience in video websites: the effect of social presence on E-loyalty[J]. International Journal of Electronic Commerce, 2018, 22 (3) : 446-476.
- [17] 叶许婕, 赵宇翔, 张轩慧. Reaction视频的用户弹幕评论行为生成机制探索: 基于认知-情感系统理论[J]. 数据分析与知识发现, 2023, 7 (2) : 1-14.
- [18] CHEN Y E, GAO Q, RAU P L P. Understanding gratifications of watching danmaku videos: videos with overlaid comments[C]//Cross-Cultural Design Methods, Practice and Impact: 7th International Conference, CCD 2015, Held as Part of HCI International 2015. Lecture Notes in Computer Science, 2015, 9180: 153-163.
- [19] 李红革, 黄家康. 弹幕文化对主流意识形态的风险挑战及其应对策略[J]. 湖南社会科学, 2022 (6) : 52-57.
- [20] 徐敬宏, 张如坤, 郭婧玉, 等. 弹幕与规训: 网络直播中的性别凝视: 基于斗鱼大数据的分析[J]. 国际新闻界, 2022, 44 (4) : 146-151, 153.
- [21] WU Q F, SANG Y S, ZHANG S, et al. Danmaku vs. forum comments: understanding user participation and knowledge sharing in online videos[C]//Proceedings of the 2018 ACM International Conference on Supporting Group Work, 2018: 209-218.
- [22] 冯钰茹, 邓小昭. 弹幕视频网站用户弹幕评论行为的影响因素研究: 以Bilibili弹幕视频网站为例[J]. 图书情报工作, 2021, 65 (17) : 110-116.
- [23] 储节旺, 李佳轩. 虚拟知识社区中用户的知识共享策略演变研究: 基于无标度网络理论[J]. 情报科学, 2023, 41 (4) : 26-34.
- [24] HE M, GE Y, CHEN E H, et al. Exploring the emerging type of comment for barrage videos: danmu[J]. ACM Transactions on the Web, 12 (1) : 1.
- [25] 王文韬, 陈千, 张肖, 等. 弹幕视角下的网络热搜健康视频关注度与情感分析[J]. 图书馆论坛, 2022, 42 (3) : 98-107.
- [26] 黄立赫, 石映昕. 面向视频弹幕的网络舆情事件监测研究[J]. 情报杂志, 2022, 41 (2) : 146-154.
- [27] IBRAHIM Z A A, SAAB M, SBEITY I. VideoToVecs: a new video representation based on deep learning techniques for video classification and clustering[J]. SN Applied Sciences, 2019, 1 (6) : 1-7.
- [28] 张腾, 倪渊, 莫同, 等. 弹幕视频的情感时间曲线聚类与传播效果[J]. 数据分析与知识发现, 2022, 6 (6) : 32-45.
- [29] YAMAMOTO T, NAKAMURA S. Leveraging viewer comments for mood classification of music video clips[C]//Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2013: 797-800.
- [30] HONG Q, WANG S, ZHAO Q, et al. Video user group classification based on barrage comments sentiment analysis and clustering algorithms[J]. Computer Engineering and Science, 2018, 40 (6) : 1125-1139.
- [31] IBRAHIM N F, WANG X J. A text analytics approach for online retailing service improvement: evidence from Twitter[J]. Decision Support Systems, 2019, 121: 37-50.
- [32] KAUR W, BALAKRISHNAN V, RANA O, et al. Liking, sharing, commenting and reacting on Facebook: user behaviors’ impact on sentiment intensity[J]. Telematics and Informatics, 2019, 39: 25-36.

- [33] HAGEN L. Content analysis of e-petitions with topic modeling: how to train and evaluate LDA models?[J]. Information Processing & Management, 2018, 54 (6) : 1292-1307.
- [34] 孙晓宁, 姚青. 多元主题场景下的用户弹幕与评论特征比较研究: 基于Bilibili网站[J]. 情报理论与实践, 2021, 44 (9) : 135-141, 121.
- [35] 刘琼, 马文婷, 范一欣. 短视频平台突发公共事件的网络情绪呈现及舆情治理: 以Bilibili网站“新冠疫情”议题为例[J]. 电子政务, 2021 (6) : 52-65.
- [36] YANG Z L, DAI Z H, YANG Y M, et al. XLNet: generalized autoregressive pretraining for language understanding[EB/OL]. [2023-03-01]. <https://arxiv.org/abs/1906.08237>.
- [37] ZHOU J L, ZHOU J, DING Y, et al. The magic of danmaku: a social interaction perspective of gift sending on live streaming platforms[J]. Electronic Commerce Research and Applications, 2019, 34: 100815.
- [38] ZOU L X, LIU X W, BUNTINE W L, et al. Citation context-based topic models: discovering cited and citing topics from full text[J]. Library Hi Tech, 2021, 39: 1063-1083.
- [39] XU L H, LIN H, PAN Y, et al. Construction of emotional vocabulary ontology[J]. Journal of Information Science, 2008, 27 (2) : 180-185.
- [40] CAO W, ZHANG K, WU H, et al. Video emotion analysis enhanced by recognizing emotion in video comments[J]. International Journal of Data Science and Analytics, 2022, 14 (2) : 175-189.
- [41] 王晓, 董迪迪, 陈思菁, 等. 社交媒体主题分布特征及其对情感倾向影响研究[J/OL]. 情报科学, 2023: 1-23[2023-07-06]. <https://kns.cnki.net/kcms/detail/22.1264.G2.20230509.1432.010.html>.

作者简介

李佳轩, 男, 硕士研究生, 研究方向: 知识管理、科技管理, E-mail: 11159892696@qq.com。
储节旺, 男, 博士, 教授, 研究方向: 知识管理。

Joint Analysis of User Attention and Sentiment Value in Online Knowledge Communities Based on Danmaku Comments

LI JiaXuan CHU JieWang
(School of Management, Anhui University, Hefei 230601, P. R. China)

Abstract: This study attempts to analyze the relationship between attention and sentiment of knowledge community users in danmaku comments, and tries to construct a coordinate to associate attention and sentiment, so as to provide a new method for the classification and promotion of knowledge videos. By obtaining a certain number of danmaku comments, we use the Sentence-LDA model and XLNet pre-training model to analyze the text of danmaku comments. We extract the subject words and keywords of the danmaku comments during the video and estimate their sentiments. Based on the user attention and sentiment to different danmaku themes, we establish a user attention-sentiment coordinate system. The results show that as the video plays, the attention and sentiment of online knowledge community users to different topics in the video change significantly. Users have different emotional evaluations of different topics in the video, and even for the same topic, users change their emotional attitude as the video progresses. Through the evolution of sentiments and themes, a new knowledge video content label can be constructed.

Keywords: Online Knowledge Community; Danmaku Comment; Joint Analysis; Knowledge Promotion; Video Classification; Sentiment Analysis

(责任编辑: 王玮)