

摘 要

随着智能设备和社交媒体的蓬勃发展,在线社交已经成为人们主要的交流方式。emoji 是一种基于图像符号的非文本语言,经常被用来表达丰富的语义和情感信息。emoji 可以弥补文本传达不了的表情动作和肢体动作等外部信息,对于理解文本中的情感信号具有重要的作用。

以往研究显示,将 emoji 信息纳入基于文本的情绪识别模型中,对于提升模型的性能起着重要作用。如何对 emoji 信息进行有效的向量表示,是在情绪识别模型中成功融入 emoji 信息的关键。现有的 emoji 向量表示方法大多采用两种方式。一种是使用基于描述文本构建 emoji 向量表示,它可以利用常规文本词嵌入资源。然而,常规文本通常难以准确传达类似于图像符号 emoji 的视觉语义,而且在不同情境下,同一个 emoji 可能具有不同的含义。另一种是基于模型训练的 emoji 向量表示,它可以提取出 emoji 所包含的语义信息。然而,传统的 emoji 词嵌入模型通常是基于单词共现规律学习向量表示,没有将 emoji 与目标情绪直接关联,得到的 emoji 向量蕴含的情绪识别信息较少。

针对上述问题,本文通过使用软标签的方式为 emoji 构建一个与情绪信息相关联的情感分布向量,软标签是预训练模型在情绪预测任务中全连接层输出的向量;再将 emoji 情感分布信息与基于预训练模型的文本语义信息相结合,提出融合 emoji 情感分布的多标签情绪识别方法(Emoji emotion distribution Information Fusion for multi-label Emotion Recognition, EIFER)。EIFER 方法的模型结构由文本语义信息模块、emoji 信息模块和多损失函数预测模块组成。文本语义信息模块使用 RoBERTa 模型抽取句子的文本语义信息,并输出语义信息向量;emoji 信息模块通过 emoji 情感分布知识库抽取出句子中 emoji 信息向量;多损失函数预测模块则融合前两个模块输出的向量合成一个新的特征向量,最后将新的特征向量输入至全连接层中进行情绪识别。EIFER 方法在使用经典的二元交叉熵损失函数的基础上,通过引入标签相关感知损失函数提升模型的情绪识别性能。在 SemEval2018 英文数据集上的情绪预测对比实验结果表明,本文提出的 EIFER 方法比已有的情绪识别方法具有更优的性能。同时,在解决具有情绪标签模糊性问题的情绪预测任务时,尤其是预测具有高度相关的情绪时,对情绪间的相关性进行建模能够有效地提升模型的预测性能。

关键词: emoji 情感分布; 多标签分类; 情绪识别; 情绪相关性

Abstract

With the booming development of smart devices and social media, online socialization has become the main way for people to communicate with each other. Emoji in text is a kind of non-textual language based on image symbols, which is often used to express rich semantic and emotion information. Emoji can make up for the external information such as expression movements and body movements that cannot be conveyed by the text, and play an important role in understanding the emotion signals in the text.

Existing studies have shown that fusion emoji information in text-based emotion recognition models plays an important role in improving model performance. How to represent emoji information effectively is the key to successfully fuse emoji information in emotion recognition models. Most of the existing emoji vector representation use two approaches. One is emoji vector representation constructed based on description text, which can utilize regular text word embedding resources, however, it is difficult for regular text to accurately express the visual semantics similar to image symbols emoji, and the meaning of the same emoji may not be the same in different scenarios. The other is the emoji vector representation based on model training, which can extract the semantic information contained in emoji, however, traditional emoji word embedding models generally learn vector representations based on the co-occurrence rule of words, without directly associating emoji with the target emotion, and the resulting emoji vector representations contain less emotion recognition information.

To address the above problems, this paper constructs an emotion distribution vector associated with the emotion information for emoji through soft labels which are the vectors output from the fully connected layer used by the pre-trained model for the emotion prediction task. Then combines the emoji emotion distribution information with text semantic information based on the pre-training model, and proposes the Emoji emotion distribution Information Fusion for multi-label Emotion Recognition method (EIFER). The model structure of the EIFER method consists of a text semantic information module, an emoji information module and a multi-loss function prediction module. The text semantic information module uses RoBERTa model to extracts the semantic information of the text in the sentence and outputs the semantic information vector; the emoji information module extracts the vectors of the emoji information in the sentence through the emoji emotion distribution knowledge base; and the multi-loss function prediction module fuses the vectors output from the first two modules and synthesize a new feature vector, and finally the new feature vectors are fed into a fully

connected layer for emotion recognition. Based on the classical binary cross-entropy loss function, EIFER method models the correlation between emotion labels by introducing label-correlation aware loss function, so as to improve the emotion recognition performance of the model. Compared experiment results of emotion prediction on the SemEval2018 English dataset have shown that the proposed EIFER method has better performance than the existing emotion recognition methods. Meanwhile, when solving the task of emotion prediction with the problem of emotion label ambiguity, especially when predicting emotions with high correlation, modeling the correlation between emotions can effectively improve the prediction performance of the model.

Keywords: emoji emotion distribution; multi-label classification; emotion recognition; emotion correlation

目 录

摘 要	I
Abstract.....	III
目 录	V
1 绪论.....	1
1.1 研究背景与意义.....	1
1.2 国内外研究现状.....	4
1.3 本文主要工作及创新点	5
1.4 本文组织架构与章节安排	6
2 相关理论及方法	8
2.1 emoji.....	8
2.1.1 emoji 起源和介绍.....	8
2.1.2 emoji 情感功能.....	9
2.1.3 emoji 情感分布学习	10
2.2 多标签情绪识别.....	11
2.2.1 多标签情绪识别的定义与流程.....	11
2.2.2 多标签情绪识别的评价指标.....	13
2.3 预训练模型	14
2.3.1 Transformer 模型.....	14
2.3.2 BERT 模型	17
2.3.3 RoBERTa 模型	18
2.4 本章小结	19
3 基于情感分布的 emoji 向量表示.....	20
3.1 现有的 emoji 向量表示方法	20
3.1.1 基于描述文本的 emoji 向量表示.....	20
3.1.2 基于模型训练的 emoji 向量表示.....	21
3.2 emoji 情感分布向量定义.....	21
3.3 基于情感分布的 emoji 向量表示.....	22
3.4 本章小结	25
4 融合 emoji 情感分布的多标签情绪识别方法	26
4.1 文本语义信息模块.....	27
4.2 emoji 信息模块	28

4.3 多损失函数预测模块.....	28
4.4 本章小结	31
5 实验与结果分析	32
5.1 实验设置	32
5.1.1 实验数据集.....	32
5.1.2 模型参数与实验环境.....	33
5.1.3 模型性能的评价指标	34
5.2 损失函数权重系数 λ 对 EIFER 模型预测性能的影响实验	35
5.3 多种模型情绪预测性能的对比实验.....	37
5.4 EIFER 模型预测情绪的相关性分析实验	40
5.5 EIFER 模型与 RoBERTa _{base} 模型对句子的预测结果分析实验	42
5.6 本章小结	43
6 总结与展望.....	44
参考文献.....	46
致 谢	51
在读期间公开发表论文（著）及科研情况.....	53

1 绪论

1.1 研究背景与意义

情绪可以简单地定义为描述一个人心理状态的具体感觉，如喜爱、愤怒、厌恶和恐惧等，换句话说，情绪是针对某事或某人的强烈感受^[1]。如今，互联网发展迅速，线上交流已经成为日常交流的主要方式，包括微博、用户评论、博客、贴吧等。用户在网络上频繁交流，分享观点，从而产生了许多有价值的文本数据，这些文本数据中蕴含丰富的情绪信息，为情绪相关的研究提供了数据支持^[2]。文本语言是人类情绪最基本、最直接的载体，在自然语言处理中，文本情绪识别因其巨大的学术和商业潜力而成为一个重要的研究课题^[3]。文本情绪识别旨在从用户社交网络的文本中提取用户的情绪状态，被广泛应用于消费者分析^[4]，企业管理^[5]，舆情监测^[6]，医疗健康^[7]等领域。

由于 emoji 本身具有丰富的情感含义，因而成为人们情感交流的一个重要媒介。emoji 作为一种非语言线索，可以比文字更生动的反映说话者的情绪状态，甚至影响整个对话的气氛。需要注意的是，在 emoji 的使用过程中，不同的性别，不同的交流平台以及交流者的文化背景等都会导致人们对同一个 emoji 的不同理解。因此，本文总结了 emoji 在个体多样性、平台多样性以及文化背景多样性中的使用差异。

(1) 个体多样性

Emoji 的使用受到性别特征和个人心理特征的影响。首先，存在明显的性别差异，尽管男性和女性对 emoji 的理解相似，但女性使用 emoji 的频率要高于男性，不过这种情况会随着交流的改变而改变^[8]。在社交场合中，女性更喜欢使用 emoji，而在私下交流中，男性更喜欢使用 emoji。在理解 emoji 方面，女性普遍觉得 emoji 更亲切、明了，并具有更深层次的含义，男性喜欢使用同一个 emoji 来加强情感的表达。当男性和女性使用同一个 emoji 与其他人交流时，其他人会觉得他们要表达的情绪是不一样的，女性适合发送一些含有深情的或具有吸引力的 emoji，而男性适合发送含有不太深情但友好的 emoji^[9]。其次，emoji 的使用也受到个人心理特征的影响。研究显示，在 Facebook 平台上，用户使用 emoji 的频率与他们性格外向和自监督特征之间存在一定的正相关性，而正面 emoji 的使用与用户的糟糕情绪之间存在一定的负相关性。根据一项以 emoji 为基础的人格测试显示，个体与 emoji 之间的相似度得分与情绪稳定性、外向性以及合群性等五大人格特质密切相关，但是与个人的责任心和开放性无关^[10]，即积极的 emoji 则与外向性格呈正相关，消极的 emoji 与情绪稳定性呈负相关。此外，emoji 的

使用也与人际关系有关^[11]。人与人之间的对话越是礼貌和疏远, emoji 就越是抽象、几何和静态。相反, 更具体、更生动的 emoji 通常被用于对话者共同对某一特定话题更有共鸣、更亲密的群体中。

(2) 平台的多样性

不同平台的风格是影响 emoji 使用的一个重要原因。emoji 在不同网络平台的结构规范中所表达的含义和在各种操作系统中的样式都会对用户使用 emoji 造成差异。虽然 emoji 使用的是 Unicode 编码, 但由于受到不同开发者的影响, 各平台系统中 emoji 的表现风格是不同的, 这有可能导致在使用过程中用户对 emoji 情感和语义的解释存在误解和分歧^[12]。此外, 研究人员对 Twitter、Facebook 和 Google 等不同的网络平台展开研究, 发现用户在不同的平台使用 emoji 时有各自的喜好。在一个平台上使用量最高的 emoji 可能在其他平台上使用非常少。例如, 相比较 Twitter 平台, 用户更喜欢在 Facebook 上频繁且主动的使用 emoji。同时, 一些研究人员关注更多边缘化的社区平台, 如 Gab。面对同一事件, Gab 用户更倾向于使用积极的 emoji 来讽刺负面信息, 而 Twitter 用户更倾向于使用 emoji 来表示怀疑^[13]。同时, 还有一些研究人员认为, 尽管在 e-mail 中不太适合使用 emoji, 但在日常交流中对于 emoji 的使用通常是保持一致的。

(3) 文化背景的多样性

Emoji 的使用是由语言、社会背景以及文化习俗的组合构成, 文化差异对 emoji 的使用产生显著影响, 而 emoji 的具体用法也与文化背景密切相关。比如, 芬兰、印度和巴基斯坦的用户会根据各自的文化背景使用特定的 emoji^[14]。根据 Hofstede 的文化维度模型, 高权力距离和放纵型国家的人更倾向于使用带有消极情绪的 emoji; 而高不确定性规避、个人主义和长期导向型国家的人则常常使用带有积极情绪的 emoji^[15]。将西班牙与中国比较, 中国用户更倾向于使用 emoji 来表达负面情绪。研究还指出, 香港和美国的用户在餐厅评论网站上使用 emoji 的方式也存在差异的, 这或许反映了潜在的文化差异^[16]。由于 emoji 使用的文化差异, 开发了一个 EmojiGrid, 用于食品相关情绪的跨文化研究, 它可靠地反映了既定的文化特征^[17]。这种差异不仅在国家之间明显, 而且在同一国家内也很明显。其次, emoji 在跨语言交流中表现出高度的语境敏感性, 这意味着它们极其依赖其语言和文本环境^[18]。例如, 研究表明, 由于英国和美国都说英语, 因此在 emoji 的使用上有很强的相似性, 但与意大利语和西班牙语进行比较时, 相似性较低。

由于 emoji 使用形式和其表达内容的多样性, 已经成为日常社交媒体中不可缺少有效工具, 全世界每天有超过 60 亿个 emoji 在移动社交媒体中被使用^[19]。2015 年, 《牛津英语词典》将 “😄” 评为年度词汇, 显示出 emoji 在网络交流

中有重大作用。由于 emoji 具有非常广阔的研究价值，许多专家学者从计算机科学、语言学、教育学和密码学等多个领域对 emoji 进行了深入的探讨：

（1）计算机科学领域

在计算机科学领域，emoji 被广泛用于对网络数据进行情感分析，与传统的文本或图片表达方式有所不同，emoji 具有独特的视觉特征和统一编码，使其成为一种独特的表达方式。一系列研究关注于 emoji 与文本、图片以及视频之间的转换^[20]。比如，Emoji2Video 提供了一种利用 emoji 搜索视频的方法^[21]。由于句子中的情感类别与用户选择 emoji 之间存在关联，Hayati 等人^[22]和 Zanzotto 等人^[23]提出了不同的方法来预测基于文本的 emoji。Kim 等人^[24]研发了一款名为 Reeboc 的系统，它能够分析聊天内容，提取不同的情绪或主题，并根据这些内容向用户推荐合适的 emoji。此外，基于 emoji 的情感分析不受语言限制，展现出跨语言的有效性，例如，Al-Azani 等人^[25]发现，emoji 也能用于分析阿拉伯语的情感。尽管一些研究指出，在正面的情感分析评论中，使用 emoji 通常会获得更高的情感分数^[26]。

（2）语言学

在语言学领域，研究的重点是 emoji 的语言功能以及它们是否有可能发展为独立语言。有学者认为，emoji 具有语义属性，既可以用作独立语言，也可以用作副语言的组成部分，为用户提供交流并成为互动的手段^[27,28]。也有一些学者认为，emoji 具有视觉修辞和文本功能，并且具有更微妙的语义，因此认为 emoji 可以独立表达含义^[29]。还有一些学者认为 emoji 不能被视为一种独立的语言，因为其含义很大程度上取决于周围的文本，只有将它们与文本相结合时，才能表达完整的语义^[30]。

（3）教育学领域

在教育领域，大部分学者关注的是 emoji 对学生学习效果的影响。例如，教师在传授知识时运用 emoji 可以使得学生更好地掌握教师所传授的知识，特别是线上教学中使用 emoji 效果更好^[31]。emoji 也可以更好的协助人们理解安全知识、处理同伴之间的关系以及了解一些情绪等比较抽象的概念，同时还可以提升人们的自我表达能力^[32]。

（4）密码学领域

在密码学领域，emoji 也有用武之地。Kraus 等人^[33]提出了 EmojiAuth 项目，旨在研究将 emoji 整合到密码中以验证移动身份认证的可用性和用户感受。相较于传统的个人识别码输入，包含 emoji 的密码更容易记，因此，基于 emoji 的身份验证可作为传统个人识别码身份验证的替代方案。

由于 emoji 在许多领域都有研究且具有一定的应用价值，特别是在计算机科学领域中的情感分析，通过深挖 emoji 的情感表达，可以对用户表达的情感进行

分析,以提升情感分析模型的性能^[34]。本文专注于探讨 emoji 在情感分析领域的研究价值,旨在建立与情绪相关联的 emoji 向量表示,并将其与预训练模型的文本语义信息相结合,用于提升模型的情绪识别性能。

1.2 国内外研究现状

Emoji 首次出现的时间是 1999 年,由日本的电信公司员工 Shigetaka Kurita 创造,最初仅由 176 个图形符号组成^[35]。emoji 是日文单词“絵(e=picture)文(mo=write)字(ji=character)”的音译,在 2010 年时,由统一码联盟(Unicode consortium)正式将 emoji 标准化,使得每一个 emoji 都有其对应的编码符号和文本解释。截止当前,emoji 已将版本更新到 15.1 版,总共有 3773 个 emoji。

Emoji 在使用过程中可以不遵循中英文语法的规则,在句子中的任何位置都可以随意添加,作用是传达说话者的主观情感,包括建立情感基调,减少话语模糊性,增强语境适宜性和加强或削弱语言行为等^[36]。由于 emoji 具有丰富的情感含义,被广泛用于表达情感,已经成为辅助情绪识别的有效手段。目前已有一些研究者提出考虑 emoji 信息的文本情绪识别模型,其中,如何对 emoji 信息进行有效表示,是在情绪识别模型中成功融入 emoji 信息的关键。经典的词嵌入模型学习 emoji 向量表示方法已经成为主流的向量表示方法,一般分为基于描述文本的 emoji 向量表示和基于模型训练的 emoji 向量表示。

基于描述文本的 emoji 向量表示方法是先用 emoji 对应的文本解释将其进行替换,之后使用特定的方法获取对应解释文本的向量表示,再通过计算得到最终的 emoji 向量表示。例如, Eisner 等人^[37]参考 word2vec 的向量学习方法,使用 emoji 的官方文本解释来学习 emoji 向量表示并构建一个 emoji2vec 词典; Singh 等人^[38]使用 emoji 的短文本描述将其替换,之后在已经训练好的词典中抽取出描述词对应的向量作为最后的 emoji 向量表示; Wijeratne 等人^[39]从 EmojiNet 中提取 emoji 语义描述文本,学习 emoji 的向量表示; Liu 等人^[40]根据人类注释的语料库训练 emoji 平滑语言模型并用于 Twitter 文本的情感分类; Kimura 等人^[41]通过分析 WordNet-Affect 中 emoji 和情感词之间的共现规律,学习 emoji 的情感向量; Aoki 等人^[42]根据情感词和 emoji 之间的关系自动生成了 emoji 向量。基于描述文本构建的 emoji 向量表示通常是利用常规的文本词嵌入资源,这些词嵌入资源都是通过大型文本库训练得到的;但是,emoji 一般以视觉方式呈现,可以传达一些视觉语义信息,常规的文本向量却通常难以表达其含义,而且在不同的情景下,同一个 emoji 可能具有不同的含义。

基于模型训练的 emoji 向量表示是将 emoji 视为一个特殊的单词,通过预训练模型进行词嵌入训练,最终生成 emoji 的向量表示。例如, Chen 等人^[43]使用 bi-sense emoji 嵌入以及使用一种基于 Attention 的 LSTM 去训练情绪识别模型,

并将 emoji 作为特殊单词输入 fastText 模型^[44]中获得 emoji 向量表示; Felbo 等人^[45]将 emoji 代表文本的情感, 并使用大量的 emoji 来预训练模型, 获得 emoji 向量表示; Yuan 等人^[46]提出一种基于 emoji 的协作注意力网络, 学习 emoji 的情感语义特征; Lou 等人^[47]使用注意力机制获得 emoji 与文本之间的关系; Wang 等人^[48]设计了一个混合情感实体识别模型, 该模型将 emoji 分为四个不同的情感类别, 然后基于该模型对情感数据进行分类。基于模型训练的 emoji 向量表示在目标数据集上对 emoji 和文本进行训练, 可以抽取出 emoji 包含的语义信息; 但是, 传统的 emoji 词嵌入模型通常是计算词与词之间的共现规律来学习向量表示, 没有将 emoji 与目标情绪直接关联, 得到的 emoji 向量表示蕴含较少的情绪识别信息。

1.3 本文主要工作及创新点

针对目前的 emoji 词向量训练过程主要是考虑训练数据中的共现规律, 没有与目标情绪直接关联, 得到的 emoji 向量蕴含情绪识别信息较少等问题。本文采用 emoji 情感分布直接建模 emoji 在各个情绪上的表达程度, 提出了一种融合 emoji 情感分布的多标签情绪识别 (Emoji emotion distribution Information Fusion for multi-label Emotion Recognition, EIFER) 方法。EIFER 方法基于软标签从训练数据中学习直接与目标情绪关联的 emoji 情感分布向量, 并构建 emoji 情感分布知识库, 再将 emoji 情感分布信息与文本语义信息相结合进行多标签情绪识别。EIFER 方法的网络构架分为文本语义信息模块、emoji 信息模块和多损失函数预测模块。文本语义信息模块采用预训练 RoBERTa 模型^[49]抽取文本包含的语义信息。emoji 信息模块借助 emoji 情感分布知识库, 抽取出句子的 emoji 情感分布向量。多损失函数预测模块将前两个模块的输出进行融合, 再通过一个全连接层对句子的情绪进行预测。EIFER 模型训练采用的多损失函数包括二元交叉熵 (Binary Cross-Entropy, BCE) 和标签相关感知 (label-correlation aware, LCA)^[50]; 其中, BCE 损失是传统的多标签分类训练目标, LCA 损失对共现标签进行建模, 可以学习情绪间的相关性。已有研究工作表明^[51], 结合 BCE 损失和 LCA 损失可以提升深度模型的情绪识别性能。最后, 在 SemEval2018 英文数据集^[52]上的对比实验表明, 本文提出的 EIFER 方法具有优于对比方法的情绪识别性能。本文的主要贡献如下:

(1) 提出一种融合 emoji 情感分布的多标签情绪识别 (Emoji emotion distribution Information Fusion for multi-label Emotion Recognition, EIFER) 方法。EIFER 方法基于软标签从训练数据中学习直接与目标情绪关联的 emoji 情感分布表示, 构建一个 emoji 情感分布知识库, 并将 emoji 情感分布信息与文本语义信息相结合进行多标签情绪识别。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/208003042066007007>