



OPEN 使用表情符号情感框架预测代码混合文本中的多标签表情符号、情感和情绪

Gopendra Vikram Singh、Soumitra Ghosh、Mauajama Firdaus、Asif Ekbal 和 Pushpak Bhattacharyya

在社交媒体时代，表情符号和代码混合语言的使用已成为在线交流的必备条件。然而，选择与代码混合文本中的特定情绪或情绪相匹配的适当表情符号可能很困难。本文提出了一项预测英语-印地语代码混合句子中多个表情符号的新任务，并提出了一个名为 SENTIMOJI 的新数据集，该数据集扩展了 SemEval 2020 Task 9 SentiMix 数据集。我们的方法基于利用情感、情绪和表情符号之间的关系来构建端到端框架。我们用简单的线性变换替换 Transformer 编码器中的自注意力子层，并使用 RMS 层范数而不是普通层范数。此外，我们采用门控线性单元和全连接层来预测表情符号并识别推文的情感和情绪。我们在 SENTIMOJI 数据集上的实验结果表明，所提出的多任务框架优于单任务框架。我们还表明，表情符号与情绪和情绪密切相关，识别情绪和情绪可以帮助准确预测最合适的表情符号。我们的工作为自然语言处理领域做出了贡献，可以帮助开发更有效的工具，用于代码混合语言中的情感分析和情感识别。代码和数据将在 <https://www.iitp.ac.in/~ai-nlp-ml/resources.html# SENTIMOJI> 上提供，以方便研究。

表情符号已经发展成为当代交流中无处不在的组成部分，为个人提供了一种在文本中传达情感和细微差别的方式。在社交媒体平台广泛采用的推动下，预测表情符号的任务在自然语言处理（NLP）领域变得至关重要。在当今的数字环境中，人们经常使用多个表情符号来表达他们的情绪和情绪，如表 1 所示。然而，辨别与特定情绪或情感相对应的精确表情符号却具有挑战性，特别是在多种语言合并的代码混合文本中。代码混合是一种在多语言社区中经常发生的语言现象，其中说话者在交流时倾向于在两种或多种语言之间切换。代码混合在情感和情感分析中引入了一层复杂性。社交媒体帖子，尤其是 Twitter 等平台上的帖子，经常混合语言，这给解读情感和丰富的情感内容带来了挑战。例如，以下推文：“多好的三柱门啊！@iammirofficial 我们爱你。Kya kar diya apney?”（翻译过来就是：“真是个小门！@iammirofficial，我们爱你。你做了什么？”）说明了这种代码混合文本，这使得识别情绪和情绪变得更具挑战性。表情符号、情绪和情感之间的联系是复杂的，对于理解代码混合文本中的含义层次至关重要。情绪（表达积极、消极或中立的观点）和情感（揭示对某个主题的感受（例如愤怒、快乐））与表情符号密切相关，正如

¹印度理工学院巴特那计算机科学与工程系，巴特那 801103，印度。印度理工学院孟买分校计算机科学与工程系，印度孟买 400076阿尔伯塔大学，加拿大阿尔伯塔省埃德蒙顿。这些作者做出了同样的贡献：Gopendra Vikram Singh 和 Soumitra Ghosh。电子邮件：gopendra.99@gmail.com

| 推文 | 英文翻译 | 表情符号 |
|---|-----------------------------|-------|
| 我爱纳伦德拉·莫迪 ab kya karun wo hein hi ache PM!!!! | 我喜欢纳伦德拉·莫迪所做的一切，他是一位好总理！！！！ | 😊, ❤️ |
| Mujhe nafrat hai 是一个应用程序，它真的很烦人。 | 我讨厌这个应用程序，它真的很烦人 | 😡, 😡 |
| Aaj kal ki media ghin ati hai 他们真的很糟糕 | 这些天媒体我感到厌恶他们真的很糟糕 | 🤢, 😡 |

表 1. 说明推文中存在多个表情符号。

以前的研究。为了说明这一点，请考虑“探索鬼屋”这条推文。尽管“探索鬼屋”最初让人感到兴奋，但厌恶呕吐表情符号的添加表明了情绪和情绪的转变。这种组合暗示着个人探索鬼屋的经历是不愉快的，并引起了厌恶的感觉。因此，这条推文整体传达了一种负面情绪，以及对鬼屋的排斥或不适感。这种情况进一步说明了表情符号、情绪和情感之间的密切关系，因为表情符号的选择可以显着改变对文本潜在情绪和情感的解释。

表情符号检测、情感分类和情绪识别等任务通常是相互依赖和关联的。当同时处理时，它们的表现会更好。预测代码混合句子中的多个表情符号的任务可以是一个利用情感、情感和表情符号之间相关性的多任务学习问题。识别给定推文的情绪和情绪可以帮助预测准确代表情绪和情绪的最合适的表情符号。同样，推文中表情符号的存在可以帮助识别情绪和情绪，从而有可能提高情绪分析和情绪识别模型的性能。

本文介绍了一项新颖的任务，该任务侧重于预测代码混合句子中的多个表情符号，同时识别用户相应的情绪和情绪。我们的方法利用零样本策略将情感信息整合到训练过程中。我们的目标是利用情感、情感和表情符号之间的相互依赖关系，创建一个全面的框架，可以准确地检测代码混合文本中的所有三个组成部分。在这项工作中，我们通过 RMS 归一化增强 F-Net，并引入门控线性单元 (GLU) 和全连接 (FC) 层来预测表情符号并识别给定推文的情感和情感。为了促进我们的研究，我们通过扩展代码混合推文的基准情绪注释 SentiMix 数据集以及多标签表情符号的手动注释来引入 SENTIMOJI 数据集。

所提出的多任务框架可以帮助弥合情感和情感分析领域的研究空白。尽管已经进行了几项研究来分别预测表情符号、情绪和情绪，但缺乏专注于在代码混合句子中同时预测它们的研究。因此，我们的工作可以通过提供对表情符号、情感和代码混合句子中情感之间关系的全面理解来推进情感和情感分析领域。先前的研究独立地研究了表情符号与情绪或情绪之间的相关性，但没有解决在多任务框架中捕获多个表情符号、情绪和情绪之间同时关系的复杂性。人类情感的复杂性以及使用多个表情符号来表达完整含义，使得通过识别推文中的多个表情符号来准确理解推文的整体情绪至关重要。因此，预测适当的表情符号对于理解推文的情绪至关重要。在处理代码混合数据时，挑战变得更加复杂，需要隐含地理解情感、情感和表情符号之间的关系，才能准确预测相关的表情符号。我们的工作引入了一种新颖的端到端多任务框架，旨在解决在代码混合数据中检测多个表情符号及其相应的情绪和情绪的挑战。据我们所知，这是首次尝试在多任务框架中利用情感和情感信息来解决代码混合数据中的多标签表情符号预测问题。

广泛的目标

这项研究的主要目标是开发一个强大的框架，用于代码混合文本中的多标签表情符号预测，并通过情感和情感分析进行丰富。我们的目标包括：

- 引入一项新颖的任务，即预测英语-印地语代码混合句子中的多个表情符号，同时识别用户相应的情绪和情绪。
- 提出零样本策略，将情感信息整合到训练过程中，促进在无监督环境下准确的情感分类。
- 开发一个端到端框架，利用情感、情感和表情符号之间的相关性来同时识别代码混合文本中的情感、情感和相应的表情符号。
- 通过对 20k 代码混合文本上的多标签表情符号进行手动注释来扩展情感注释的 SentiMix 2020 数据集，从而生成 SENTIMOJI 数据集。
- 通过在 SENTIMOJI 数据集上超越最先进的模型来证明所提出的框架的有效性。
- 开源代码和数据以促进进一步研究 (<https://www.iitp.ac.in/~ai-nlp-ml/resources.html#SENTIMOJI>) 。

我们论文的结构如下：在“相关工作”部分，我们对与我们的研究相关的先前研究进行了全面的回顾。接下来，我们在“方法论”部分介绍了我们提出的多任务实验方法，其中概述了我们开发模型所采取的步骤。我们在“关于数据集”部分详细描述了数据集及其注释，其中我们提供了有关实验中使用的语料库的信息。“实验、结果和分析”部分介绍了进行的实验、获得的结果以及我们对这些结果的分析。最后，在“结论”部分，我们总结了我们的发现并提出了未来研究的潜在途径。

相关工作

由于表情符号在社交媒体中的广泛使用，预测表情符号已成为 NLP 中的一项关键任务。因此，一些研究集中在表情符号、情绪和情绪的分析上，这些都是 NLP 领域的重要领域。

Cowie 和 Cornelius 是探索情感在自然语言处理 (NLP) 中重要性的先驱研究人员之一。随后的研究在 NLP 中使用了表情符号，包括 Eisner 等人。 ，他引入了预先训练的 emoji2vec 嵌入，可以分析表情符号的用法以及不同表情符号之间的关系。其他研究，例如 Felbo 等人。 ，周等人。和金等人。开发了模型来预测文本中的表情符号，而 Barbieri 等人和王等人则使用预先训练的 emoji2vec 嵌入分别创建用于单标签和多语言表情符号预测的模型。Guibon等人提出了一种用于即时通讯的表情符号推荐模型，而Wijeratne等人则开发了一种表情符号意义消歧模型。此外，Santhanam 等人还使用了表情符号。和胡辛等人。分别了解危机事件并对情绪进行分类。Al-Baaj 等人也将表情符号用于情感分析。和陈等人。 ，以及周等人。用于产生情绪反应。此外，马等人。提出了一种基于变压器的网络，用于多标签表情符号预测，该网络在基准数据集上优于以前的模型。这些研究证明了表情符号在 NLP 中的重要性，并为该领域的研究提出了新的方向。卡达等人。提出了一种新颖的架构，用于生成富含表情符号的讽刺句子。当仅靠单词可能无法有效传达讽刺时，它利用效价反转和语义不一致来创建讽刺句子。

近年来，代码混合文本中的情感和情绪分析引起了研究人员越来越多的兴趣。已经提出了各种技术来提高代码混合数据中情感和情绪检测的准确性。例如，Chakravarthi 等人。开发了一个德拉威代码混合数据集，用于识别泰米尔英语、卡纳达英语和马拉雅拉姆英语文本中的情感和攻击性语言。Mohbey 等人的研究。采用卷积神经网络 (CNN) 和长短期记忆网络 (LSTM) 相结合的混合方法来有效分析猴痘推文中的情绪。这项工作建立在推文情感分析研究奠定的基础上，结合了 CNN 和 LSTM 的优势来改进情感分类。Yadav 等人建议使用预先训练的语言模型将知识从单语文本转移到代码混合文本以进行情感分析。张等人代表了人类情感中情感与情绪之间的紧密联系，导致了NLP中情感分析和情感识别的相互依赖。为了解决会话上下文依赖、多模态交互和多任务相关性，所提出的 M3GAT（一种多模态、多任务交互式图注意网络）在统一框架中解决了这些挑战。相比之下，Wang 等人提出了一种联合因子图模型，该模型考虑了不同语言和情感之间的相互作用，用于代码混合文本中的情感预测。王等人。引入了双语注意力网络（BAN）模型，该模型利用注意力机制来聚合单语和双语信息词，以预测代码混合数据中的情绪。肖等人。引入了用于基于多模态方面的情感分析的网络，对齐和融合来自不同模态的信息，解决了以前主要关注文本数据的情感分析的局限性。毛等人。探讨了训练数据中的偏差如何影响这些任务中语言模型的性能，并深入了解有偏差的预测的局限性和潜在的缓解措施。现有的研究主要集中在单标签表情符号分类任务上，在理解多标签表情符号识别所需的挑战和策略方面存在差距，而 Twitter 数据的复杂多样的背景试图填补这一空白。他等人提出了一种利用元学习技术来提高标签数据稀缺场景下基于方面的情感分析的有效性和效率的方法。Mao等人引入了一种新颖的门控机制，以促进多任务学习塔之间的通信，增强模型同时执行基于方面的情感分析和顺序隐喻识别的能力。Meena 等人进行的研究引入了一种混合深度学习方法来检测 Monkeypox 推文中的情绪极性。此外，该研究还结合了知识图表示技术，展示了结合情感分析和知识图方法论的跨学科方法。最近的研究证明了先进的机器学习技术在表情符号预测任务中的潜力。研究旨在探索表情符号、情绪和情绪之间的相关性。诸如此类的研究已经显示出使用这些技术的有希望的结果。后一项研究进一步结合了多模态信息（包括文本和图像）来进行多模态表情符号预测。其他研究也应用了表情符号预测，如图所示。例如，使用表情符号进行反讽检测，同时利用表情符号信息来识别辱骂性语言。多项研究提出了提高表情符号预测准确性的新技术和模型。引入了标签式注意机制，以更好地理解表情符号预测任务中的细微差别，同时结合表情符号信息进行推文分类。一些研究旨在使用屏蔽语言模型实现任务的零样本学习。张等人。探索了一种多任务学习方法，该方法侧重于对话数据中多模式讽刺检测和情感分析的共同和独特方面。通过考虑共享和不同的特征，所提出的模型旨在提高在不同背景下这两项任务的性能

方式。张等人提出了一种多任务学习模型，本文解决了会话数据中多模态讽刺、情绪和情感识别的挑战。情感信息也被用来提高表情符号预测的准确性，如先前的调查所示，例如 Ilyas 等人的工作。 ，专注于代码混合的罗马乌尔都语和英语文本中的情绪检测。通过利用针对代码混合数据定制的深度学习技术，这项研究通过解决情感检测中代码混合带来的独特语言挑战，为更广泛的情感分析领域做出了贡献。在 中，表情符号被用于识别攻击性语言的框架中。此外，研究表明，支持向量机 (SVM) 模型在表情符号预测任务中优于递归神经网络 (RNN) 框架。其他研究探索了用于表情符号预测和识别的各种机器学习技术。例如，用于表情符号预测的梯度提升回归树方法 (GBM) 和双向 LSTM，同时研究用于表情符号识别的朴素贝叶斯和 RNN。利用表情符号中的情感知识来识别推文中的情感。此外，一些研究提出了新颖的框架，例如在 中用于英语表情符号预测的带有注意框架的残差 CNN-LSTM，以及在 中用于精确表情符号预测的基于向量相似性的方法。最后，在 中引入了用于情感分析和表情符号预测的多语言数据集印地语、孟加拉语和泰卢固语，强调需要探索表情符号在不同语言和文化中的使用和预测。总的来说，这些研究强调了使用先进的机器学习技术和多模态信息来理解和预测表情符号在不同环境下的使用的重要性和潜力。

最近，无监督学习在 NLP 社区中越来越受欢迎，在各种任务的零样本设置中显示出了有希望的结果。在情感识别中，小样本学习已被用于将知识从现有数据集转移到看不见的数据。针对情感分析，提出了一种两阶段情感检测方法，在第一阶段使用无监督零样本学习模型计算34种情感的概率。然后，这些概率用于在第二阶段训练监督机器学习分类器，使用集成学习来预测情感标签。通过利用表情符号的语义通用性，视觉、文本和多模式表情符号预测的价值已在 MSCOCO 数据集上得到了证明。此外，为了克服视觉情感分析中小数据集的局限性，作者建议从社交媒体中发现的丰富数据中学习简洁的图像嵌入。

表 2 简要概述了相关工作部分中按各自主题分类的相关研究。在本研究中，我们的目标是通过开发一个多任务框架来区别于之前的研究，该框架结合了情感和情感信息来检测多标签表情符号、对情感进行分类并分析代码混合文本中的情感。据我们所知，这是在代码混合文本的多任务框架中同时解决所有三个任务的首次尝试。我们相信，与之前专注于个人任务的研究相比，我们的方法将提供更准确和更全面的结果。

方法论

在本节中，我们讨论提出的代码混合 RMS 傅里叶变换器 (CM-RFT) 框架，用于预测代码混合输入上的多标签表情符号、情绪和情感。

| 话题 | 学习 | 概括 |
|--------|-----------|--|
| 表情符号预测 | 艾斯纳等人。 | 引入了预先训练的 emoji2vec 嵌入，可以分析表情符号的用法以及不同表情符号之间的关系。 |
| | 费尔博等人。 | 开发模型来预测文本中的表情符号，增进对自然语言中表情符号用法的理解。 |
| | 周等人。 | 创建单标签表情符号预测模型，提高表情符号预测任务的准确性。 |
| | 巴比耶里等人。 | 利用预训练的 emoji2vec 嵌入创建单标签表情预测模型，增强表情预测能力。 |
| | 王等人。 | 使用预先训练的 emoji2vec 嵌入开发多语言表情符号预测模型，将表情符号预测的范围扩展到多种语言。 |
| 情感分析 | 阿尔-巴杰等人。 | 探索表情符号在情感分析中的使用，强调表情符号在文本数据中传达情感方面的作用。 |
| | 陈等人。 | 研究了表情符号与社交媒体数据中的情绪之间的关系，为情绪分析技术提供了见解。 |
| | 毛等人。 | 引入了一种新颖的门控机制，用于基于方面的情感分析和顺序隐喻识别。 |
| | 桑塔南等人。 | 利用表情符号解读危机事件，展示表情符号在紧急情况下情绪分析的应用。 |
| | 莫贝等人。 | 采用混合 CNN-LSTM 方法对 Monkeypox 推文进行情感分析，提高社交媒体数据中情感检测的准确性。 |
| 情绪检测 | 张等人。 | 提出了 M3GAT，一种多模式、多任务交互式图注意力网络，解决了 NLP 中情感分析和情感识别的相互依赖关系，为分析文本数据中的情感和情感提供了统一的框架。 |
| | 伊利亚斯等人。 | 专注于代码混合的罗马乌尔都语和英语文本中的情绪检测，强调在多语言数据中检测情绪的挑战和机遇。 |
| | 张等人。 | 探索了用于多模态讽刺检测和情感分析的多任务学习方法，证明了在情感分析任务中结合多种模态的有效性。 |
| 多模态分析 | 肖等人。 | 开发了基于多模态方面的情感分析网络，整合文本和视觉信息以提高情感分析的准确性。 |
| | 张等人。 | 提出了一种用于多模式讽刺、情感和情感识别的多任务学习模型。 |
| | 伊利亚斯等人。 | 利用少样本学习来检测代码混合文本中的情绪。 |
| | 张等人。 | 研究了 NLP 任务中的无监督学习，在各种任务的零样本设置中显示出有希望的结果。 |
| 零样本学习 | 特斯法格吉什等人。 | 提出了一种使用无监督零样本学习的两阶段情绪检测方法，展示了零样本学习在情绪检测任务中的潜力。 |

表 2. 按各自主题分类的相关研究的简要概述。

在带有表情符号和情感标记示例的英语-印地语代码混合句子的背景下，主要目标是设计一个全面的多任务框架，该框架应该预测句子中的多个表情符号，同时识别用户表达的相应情感和情绪。值得注意的是，情绪检测任务以零样本的方式进行，利用从其他任务获取的信息而无需直接训练。

设 x 代表输入句子， y 表示相关的情感标签， y_e 表示相应的情感
替换标签情感以及的集合封装在句子中。多任务框架的目标函数定义如下：

$$\min \frac{1}{N} \sum_{i=1}^N L(f(x; \theta), y) + L(f(x; \theta), y) + L(f(x; \theta), y)$$

(1)

这里， f ， f_e 和 f_s 分别表示表情符号、情感和情绪任务的预测函数。
是模型参数，而 λ 和 λ_e 作为正则化系数，确保情感和情绪任务在整体目标函数中的均衡贡献。最后， L ， L_e ，和 L_s
是表情符号、情绪和情感任务各自的损失函数。
为了说明上下文，请考虑表 3 中的示例，其中我们有一个代码混合句子，表达了对当前政治状况的负面情绪。用户表达的情绪是一种失望。多任务框架旨在预测此类句子中相应的表情符号、情绪和情感。
情绪检测任务以零样本的方式解决，这意味着它不接受直接训练，而是依赖于从其他任务中学到的信息。在这种情况下，句子的情绪（负面的）可以告知情绪检测任务用户可能表达的情绪。

道德声明

我们使用公开数据集创建了我的资源。我们遵守数据使用准则，没有侵犯任何版权问题。我们的研究没有涉及人类参与者。我们的机构审查委员会也审查并批准了这项研究。我们将通过适当的数据协议机制使代码和数据可用于研究目的。

拟议框架

本节描述了所提出的代码混合文本情感分析方法中涉及的不同组件。图 1 展示了我们方法的整体架构。

SentencePiece 分词器

在这项工作中，我们使用 SentencePiece 对推文进行标记。该工具将推文视为 Unicode 字母序列，并采用字节对编码 (BPE) 和一元语言模型将句子转换为子词。 BPE技术首先将数据集中的所有字符都包含在词汇表中，并逐渐获得一组合并规则。在训练阶段，一元语言模型通过以概率方式选择子词来生成多个子词分段。输出是一组子词序列，这降低了输入的维度，特别是对于罕见或未见过的单词。

代码混合嵌入生成

由于词汇外 (OOV) 单词，预训练的单词嵌入可能在代码混合数据上表现不佳。为了解决这个问题，我们在可用的代码混合语料库本身上训练了词嵌入。然而，为代码混合数据选择最合适的嵌入模型可能具有挑战性。因此，我们尝试了三种不同的嵌入模型，并将它们的输出连接起来以获得更好的性能。

字符级词嵌入

为了提取代码混合数据的字符级特征，我们遵循字符级词嵌入的工作并使用字符级词嵌入。由于基于 RNN 的编码的训练计算成本较高，并且性能并不明显优于卷积神经网络 (CNN)，因此我们利用 CNN，后跟最大池化层，以简化训练。

| 印度英语输入 | 预测 |
|---|------------|
| Aaj kal ki raajneeti bahut kaminee ho gayi hai。感到失望 | |
| 就目前的情况而言 | 情感标签: 悲伤 |
| | 情绪标签: 负面 |
| 中文翻译: 如今，政治变得非常不诚实。 | 预测的表情符号: 😞 |
| 对目前的状况感到失望。 | |

表 3. 示例数据点和拟议框架的预期输出的图示。

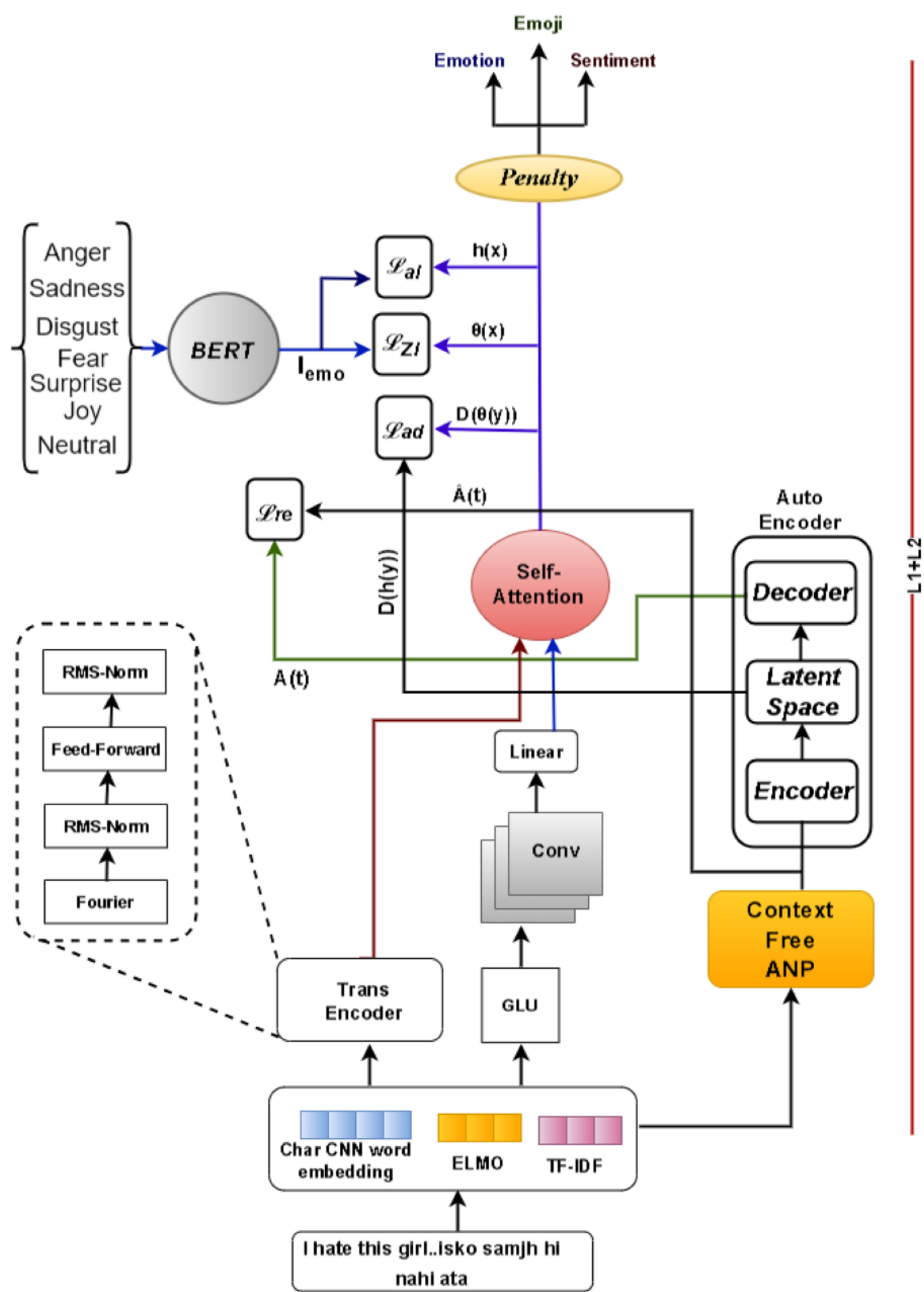


图 1. 我们提出的代码混合 RMS 傅里叶变换器 (CM-RFT) 框架的架构图。

上下文级别的词嵌入
对于上下文表示，使用 ELMO，它将每个标记表示为依赖于整个句子的向量，允许单词根据上下文具有多种含义。单词的 TF（词频）和 IDF（逆文档频率）（TF-IDF）的乘积是单词在文档中的频率乘以它在文本语料库中的唯一性的乘积。TF-IDF 文档频率模型有助于降低语料库中频繁出现的单词的权重。该方法侧重于语料库中的独特短语而不是重复的单词，从而产生更准确的模型。

代码混合 RMS 傅里叶变换器 (CM-RFT)

在本研究中，我们提出了代码混合 RMS 傅立叶变换（CM-RFT）模型，该模型结合了傅立叶变换模型和带有卷积层的门控线性单元（GLU），以更少的开销增加更多的注意力。在编码器层中，采用傅里叶变换器代替自注意力机制。它使用一维傅里叶变换来变换输入数据的序列维度和隐藏维度。结果是一个复数，可以表示为实数乘以虚数单位（I），仅存储结果的实数，避免需要修改（非线性）前馈子层或输出层以适应复杂的值。

编码的输出序列经历残差连接，然后是 dropout 和归一化，之后通过位置前馈层，然后是 dropout、残差连接和 RMS 层归一化。RMS 层范数是 LayerNorm 的简化版本，它省略了均值中心化过程，并使用均方根 (RMS) 统计量标准化层激活，如等式 1 所示。(1).

均方根 (a)
$$\frac{1}{n} \sqrt{\sum_{i=1}^n a_i^2}$$
 (2)

位置前馈层将数据从隐藏维度转换为 pf ，这通常是很多 h 表示。该技术基本上在神经网络前使用了逼近激活函数和 dropout。将卷积应用于序列是一个逻辑概念，可以用更少的开销添加额外的注意力。我们用由线性层和深度卷积组成的更轻的变体代替标准卷积，以进一步最小化计算。

自我关注
接下来，我们在 Trans-encoder 输出和 GLU 输出之间应用自注意力（SA），使模型能够有效地结合从两个信息源学到的特征。其背后的想法是，它可以增强模型捕获输入数据中复杂关系和模式的能力。

自动编码器
为了更好地理解和分析文本数据中的情感，我们利用了一种称为 Context-Free Grammar-Noun-Adjective-Pairs (Context Free ANP) 的技术。这种方法从句子中提取成对的形容词和名词来捕获文本的含义。我们使用 GitHub (https://github.com/StatguyUser/Context_Free_Grammar-Noun_Adjective_Pairs) 上现有的该技术实现来从文本数据中提取 ANP 对。为了学习 ANP 对的潜在表示，我们采用自动编码器，将 ANP 特征作为输入并学习它们在潜在空间中的压缩表示。所提出的方法利用文本和类语义知识来生成文本中传达的情感的更精确的表示。
为了训练自动编码器，使用了两种不同的损失函数。第一个损失函数是Alignment损失，其目的是最小化提取的 ANP特征和预测的ANP特征之间的差异。第二个损失函数是对抗性损失，它减少了编码的类特征与潜在空间中相应类质心之间的差距。对齐损失有助于确保从文本中提取的 ANP 对在潜在空间中准确表示。另一方面，对抗性损失鼓励编码的类特征更接近相应的类质心，这有助于提高情感分析的准确性。对齐损失和对抗损失的数学表示在本文的“训练和推理”部分中进行了更详细的讨论。

惩罚
我们加入惩罚值来改进代币的预测。目的是增强模型对不同标签和输入帖子之间联系的理解。在损失函数中添加惩罚的动机是在信息提取任务中为标记标记定义决策边界的挑战。这种模糊性使得基本的 softmax/sigmoid 分类器很难准确地地区分类别，并可能导致某些样本的错误分类。下面的等式最初代表了 softmax 和 sigmoid：

$$L = - \frac{1}{b} \sum_{i=1}^b \log \frac{\exp}{\sum \exp} \quad (3)$$

$$L = - \frac{1}{b} \sum_{i=1}^b \frac{1 \text{ 经验}}{\quad} \quad (4)$$

在哪里 $l \in \mathbb{R}$ 代表的特征是 $i \ b$ 代表批量大小, bb 代表偏差, and $W \in \mathbb{R}$ 表示权重矩阵。
为了解决信息提取任务中标记标记的决策边界难以确定的问题，可以使用Insightface损失技术对特征进行归一化并根据它们之间的角度差评估特征相似性。通过向角度添加惩罚值 x 来更新损失函数 θ ，这有助于更快地收敛特征。 softmax 和 sigmoid 的损失函数更新如下：

$$L_{\text{软最大}} = - \frac{1}{b} \sum_{i=1}^b \log \frac{\exp(\theta_i)}{\sum_{j=1}^N \exp(\theta_j)} \tag{5}$$

$$L = - \frac{1}{b} \sum_{i=1}^b \frac{1}{\text{经验} + \text{经验}} \tag{6}$$

在给定的上下文中， θ 表示重量之间的角度， W 和特点 l 是放大器函数。
等式 \exp 用于计算正样本的相似度得分，而 \exp 用于计算负样本的相似度得分。惩罚值 x 在分类边界上增加了一个余量，以提高特征的收敛速度。

训练和推理

在本节中，我们概述了训练模型的过程，并解释了如何预测情绪。
我们的模型使用四种不同的损失函数以端到端的方式进行训练。

重建损失

我们的目标是通过使用自动编码器来对齐学习的潜在空间中的标签特征和形容词-名词对特征的结构。自动编码器将重建形容词-名词对特征并生成潜在特征，同时保留情感相关信息。自动编码器参数的优化是通过减少损失函数来执行的，损失函数测量自动编码器的输入和输出之间的相似性。

$$L = || \hat{A}(t) - A(t) ||$$

自动编码器的输入和输出嵌入特征由 \hat{A} 表示 A 和 A 分别。

对准损失

我们的目标是对齐自动编码器的潜在空间和标签语义空间，以便产生的标签表示与潜在情感概念更相关。这是通过优化以下损失函数来实现的：

$$L = ||h(x) - \phi(l)||$$

基本情感定义。自动编码器生成的潜在空间嵌入，并且 l 是情感类别。
目标函数是通过对齐 $h(x)$ 和 $\phi(l)$ 来实现的。
损失和重建损失相结合来实现的：
 $L = L_{\text{align}} + L_{\text{recon}}$

情感特征

我们采用预训练的 BERT（基础）模型来编码 Ekman 基本情感类别（愤怒、厌恶、悲伤、喜悦、惊讶、恐惧、恐惧）的语义特征信息。此外，我们认为中性类可以容纳不属于 Ekman 分类范围的实例。从 BERT 获取特征消除了进一步人工注释的需要。

零镜头损失

我们模型的目标是最小化由以下形式表示的文本组合特征之间的差异 $\theta(x)$ ，以及标签的语义特征，表示为 $\phi(l)$ ，通过优化。

$$L = ||\theta(SA(x)) - \phi(l)||$$

对抗性损失

我们的目标是最小化文本判别能力之间的距离（ $\theta(x)$ 代表 $SA(t)$ ）以及特征中包含的丰富的情感结构数据 $\phi(l)$ 。这是通过使用一个判别器 D 使得 A 的输出特征 $\theta(x)$ are 旨在愚弄鉴别器网络的对抗性限制，尽可能与 ANP 特征没有区别。

$$L = E(\log D(h(y))) - E(\log D(\theta(y)))$$

在哪里 $\theta(y)$ 定义文本的特征， $h(y)$ 定义潜在特征空间。

联合损失

我们的模型是使用前面描述的各种损失函数的组合来训练的。

$$L = L_{\text{align}} + L_{\text{recon}} + (L_{\text{zero}} + L_{\text{adv}})$$

标签预测

当提供由文本和包含情感的标签集合组成的推文测试集时，我们的模型将通过简单的最近邻（NN）搜索来执行标签分类。测试推文和标签被输入到嵌入中以生成

$$\theta(t) \text{ and } \phi(l)。$$

$$||\theta(t) - \phi(l)||$$

模型训练

我们通过统一的损失函数来训练我们的模型，如式（1）所示。（7）。

$$L = L + L + L \tag{7}$$

L 是表情符号任务的损失函数，可以定义为每个预测表情符号的二元交叉熵损失之和：

$$L = - \sum_{j=1}^k y \log(y) + (1 - y) \log(1 - y) \tag{8}$$

其中 k 是可能的表情符号的数量， y 是第 i 个句子中第 j 个表情符号的二进制标签，并且 \hat{y} 是第 j 个表情符号的预测概率。
 L 是情感任务的损失函数，可以定义为预测情感的负对数似然：

$$L = - \text{对数} \frac{\exp(f(x;\theta)y)}{\sum \exp(f(x;\theta))} \tag{9}$$

其中 c 是可能的情绪数量， $f(x;\theta)$ 是第 j 个情绪的预测分数。
 L 是情感任务的损失函数，可以定义为二元交叉熵损失：

$$L = -y \log(\hat{y}) - (1 - y) \log(1 - \hat{y}) \tag{10}$$

在哪里 y 是第 i 个句子的情感二元标签，并且 \hat{y} 是预测的阳性概率情绪。

关于数据集

这项研究基于 SentiMix 数据集（可在 <https://zenodo.org/record/3974927#.ZEva4pFBxH6> 获取），该数据集最初在 SemEval 2020 共享任务的任务 9 中引入。该数据集包含大约 20,000 条推文，分为训练集、测试集和开发集，每个集都用一个情绪标签（积极、消极或中性）进行注释。
为了促进对代码混合文本中多个表情符号、情绪和情感的相互关联性的调查，我们引入了 SENTIment 和 eMOJI (SENTIMOJI) 数据集。情感任务在训练集、测试集和开发集上的数据分布如图 2 所示。此外，有关数据集的详细统计数据可以在表 4 中找到。

SENTIMOJI 数据集的主要亮点是：

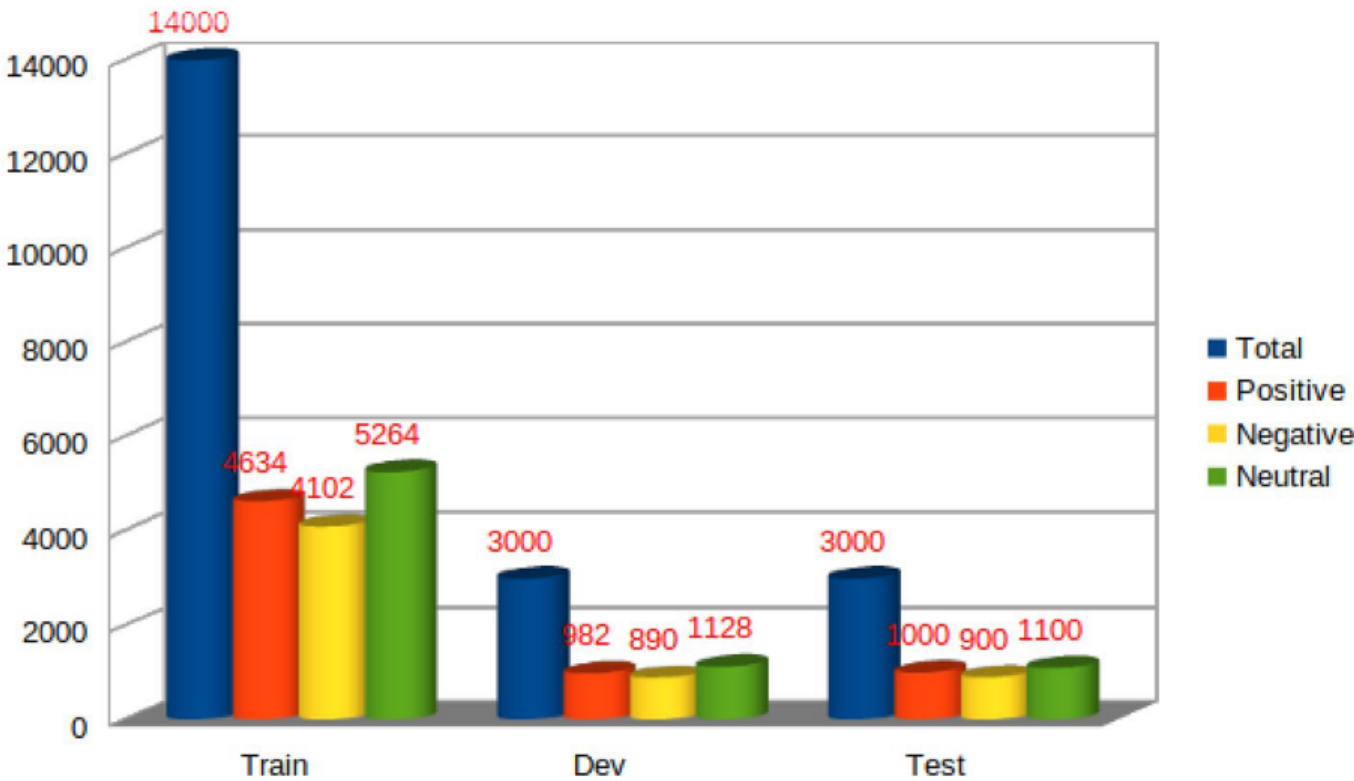


图 2. SENTIMOJI 数据集中的实例在情绪类别上的分布。

| 统计数据 | 火车 | 有效的 | Test |
|---------------|--------|-------|-------|
| # CMI（代码混合索引） | 25.32 | 25.53 | 25.13 |
| # 话语数 | 14,000 | 3,000 | 3,000 |
| 正类数量 | 4,634 | 982 | 1,000 |
| 中性类的数量 | 5264 | 1,128 | 1,100 |
| 负类的数量 | 4102 | 890 | 900 |
| 每个话语的平均表情符号数量 | 4 | 3 | 3 |
| # 唯一令牌 | 7781 | 1189 | 2458 |

表 4. Mu-Emoji 数据集的数据集统计。

- SENTIMOJI 数据集经过精心设计，可以适应代码混合语言的复杂性。
- 它包括表情符号的多标签注释，提供对情感表达更丰富的理解。
- 注释包含情绪和情感信息，提供用户表达的整体视图。
- 数据集跨越多个平台，保证了数据源的多样性。

数据标注

表情符号注释

为每个文本注释多个表情符号的过程涉及三个注释者。每个注释者的任务是用最能代表所表达的情感的表情符号来标记给定的文本。为了确保一致性，为注释者提供了一组预定义的表情符号，特别是 中提到的表情符号集中的前 64 个表情符号。这些表情符号充当注释者在标记文本时可以从中进行选择的池。注释过程是独立进行的，注释者禁止相互交流，以避免偏见或影响。 SentiMix 数据集中的每个实例（包括训练集、测试集和验证集的样本）都会提供给注释者进行标记。然后比较他们的注释以评估一致性程度。如果出现差异，注释者之间进行讨论或与高级注释者协商来解决分歧。最终的标签是根据注释者达成的共识确定的。总体而言，注释过程取得了令人印象深刻的 0.75 的 Fleiss-Kappa 分数，表明了高度的可靠性。值得注意的是，用于标记 SentiMix 数据集的所选表情符号 (, , , , , , , , , , , , , ,) 涵盖了各种情感 and 情绪。



情感注释

在处理情绪检测任务时，我们采用了零样本方法，无需专门标记的情绪数据即可进行模型训练。然而，为了评估我们模型的有效性，我们使用基于 Ekman 六种基本情绪的情绪类手动注释了 SentiMix 测试集中的每个实例：愤怒、厌恶、恐惧、快乐、悲伤和惊讶。此外，我们还添加了一个中性类来容纳不完全符合 Ekman 分类方案的实例。情绪检测的标注过程由标注者独立进行，并使用 Fleiss-Kappa 评分来衡量标注者之间的一致性，得到了值得称赞的 0.78 分，表明标注质量很高。表 5 中列出了来自 SENTIMOJI 测试集的示例注释实例，说明了指定的情绪类别。

数据集验证

SENTIMOJI 数据集的有效性是通过利用表情符号矩阵进行细致的验证过程来确保的。该矩阵交叉引用注释以减少任何差异，从而提高数据集用于后续分析和模型训练的精度和可靠性。 Emoji 矩阵是一个定义不同表情符号语义的矩阵，其灵感来自于 WordNet 的词汇矩阵。它由 14 列组成，每列

| 发声 | 英文翻译 | 表情符号1 | 表情符号2 | 情感 | 情绪 |
|--|--|-------|-------|----|----|
| @Reema sahanii suar tere jaiso ko sikhaya ja hi nahi sakta 因为你是suar hai | @Reema sahanii 像你这样的猪人永远学不到任何东西 因为你是猪 | | | | |
| @ Payal _ Rohatgi @ YouTube 推文人们在什么时候做 它们与 ki koi ghaas nahi dalta 点无关 | @ Payal _ Rohatgi @ YouTube 推文人们在什么时候做 它们无关紧要，没有人在乎你 | | | | |
| @redribbonmusik Zindagi har lamha hai khushi kar isse dosti o yaara ve 我喜欢这个 yaara | @redribbonmusik *随机歌曲* 我喜欢这个 | | | | |
| 非常感谢你 apka ek 新印度 banane ke liye Koti koti pradam apko | 非常感谢你创造了一个新印度 | | | | |

表 5. 测试 SENTIMOJI 数据集中的示例实例。

代表该任务所考虑的不同表情符号，以及代表与该表情符号关联的不同关键字的行。表 6 显示了不同的表情符号及其相应的关键字。

矩阵中的每个单元格的值为 1 或 0，具体取决于表情符号的描述是否包含关联的关键字。更具体地说，如果表情符号的描述包含特定关键字，则为矩阵中的相应单元格分配值 1，如果不包含，则为其分配值 0。

表情符号矩阵的创建方式如下：其行代表关键字集 (k_1, k_2, \dots, k_n)，列代表不同的表情符号 (e_1, e_2, \dots, e_m)，如表所示6. 矩阵中每个单元格的值由下式确定：

$$Y = \begin{cases} 1, & \text{如果 } e \text{ 的描述包含 } k \\ 0, & \text{否则} \end{cases}$$

(11)

该矩阵有助于分别识别表情符号和关键字的多义性和同义词性质。一词多义是指一个表情符号可以有多个关键字来表达其含义，而同义词意味着一个特定的关键字可以用来表示多个表情符号。

为了评估我们提出的框架的有效性，我们使用等效组代替原始表情符号，并用表情符号矩阵确定的相应组替换预测的表情符号。这构成了我们的最终预测。通过验证表情符号矩阵，我们确保了注释数据集的精度，这让我们有信心使用这些数据进行后续分析。

与现有数据集的比较

我们将我们提出的数据集 SENTIMOJI 与现有的表情符号预测数据集进行比较。表 7 总结了比较。中提出的数据集包含英语推文的单标签和多标签表情符号注释，而提出了两种语言的单标签表情符号数据集。最近，有人提出了针对英语推文的表情符号和情感注释数据集。与这些现有数据集相比，SENTIMOJI 是独一无二的，因为它本质上是代码混合的，包含英语-印地语代码混合的推文。此外，SENTIMOJI 不仅包含多标签表情符号注释，还包含情绪和情感信息，使其成为表情符号预测更全面的数据集。包含情绪和情感信息至关重要，因为表情符号通常用于传达消息的这些方面。

| 表情符号 | 意义 | 关键词 |
|------|-----------|--------------------------|
| 😡 | 红脸 | 愤怒、受伤、愤怒、疯狂 |
| 😐 | 中性脸 | 轻微的恼怒和担忧或面无表情的幽默感 |
| 😘 | 脸上飞吻 | 吻，飞吻，心，爱，嘴唇 |
| 😞 | 沉思的脸 | 失望、受伤或孤独 |
| 😡 | 撇嘴脸 | 愤怒、受伤、愤怒、疯狂 |
| 😓 | 失望的 | 失望、悲伤、压力、遗憾和悔恨 |
| 😭 | 悲伤但如释重负的脸 | 悲伤，如释重负（事情并没有那么糟糕） |
| 🤢 | 呕吐表情符号 | 身体疾病、厌恶 |
| 🤮 | 恶心的脸 | 身体疾病、厌恶 |
| 😭 | 脸上满是泪水 | 泪水、喜悦、幸福 |
| ❤️ | 红心 | 心、爱、激情、浪漫、强烈、欲望、死亡、邪恶、寒冷 |
| 😭 | 在地板上打滚 | 泪水、喜悦、幸福 |

表 6. 表情符号 SENTIMOJI 数据集的不同键。

| 数据集/论文 | 混合代码 | 情绪 | 表情符号 | |
|-------------------------------------|------|----|------|-----|
| | | | 单身的 | 多标签 |
| 用于表情符号预测的多分辨率注释 | ✗ | ✗ | ✓ | ✓ |
| 深莫吉 | ✗ | ✗ | ✓ | ✓ |
| SemEval 2018 任务 2: 多语言表情符号预测 | ✗ | ✗ | ✓ | ✗ |
| 表情符号可以预测吗? | ✗ | ✗ | ✓ | ✗ |
| 社交媒体上情绪的表情化: 纵向 Twitter 情绪数据集的收集和分析 | ✗ | ✓ | ✓ | ✗ |
| SENTIMOJI (我们的) | ✓ | ✓ | ✓ | ✓ |

表 7. 与现有数据集的比较。

数据集意义

SENTIMOJI 数据集的重要性在于它能够克服现有资源的限制。值得注意的是，SENTIMOJI 作为一个独特的代码混合数据集脱颖而出，提供多标签表情符号注释以及情绪和情感信息。如表 7 所示，我们的数据集拥有全面的统计数据，展示了其广泛且多样化的表情符号，使其成为迄今为止已知的最大的多标签表情符号数据集。通过采用代码混合语言并为表情符号以及情绪和情感数据提供多标签注释，SENTIMOJI 为更细致的分析和模型开发铺平了道路。对于努力全面了解代码混合环境中的情感表达的研究人员和从业者来说，该数据集已成为不可或缺的资产。

实验装置

我们的研究重点是一个多任务框架，该框架解决三个不同的任务：表情符号检测、情绪检测和情绪检测。我们为每个任务定义实验设置如下。

- 表情符号分类 (mE)
 - SENTIMOJI 数据集包含 12 个不同的表情符号，每条推文可以与多个表情符号相关联。
- 情感分类 (S):
 - 数据集中的每条推文都与三个情绪类别之一相关联：积极、中性或消极。
 - 我们利用 one-hot 向量来表示与每条推文相对应的情感类别。
- 情绪分类 (E):
 - 数据集中的每条推文都与五种情绪类别之一相关联：愤怒、悲伤、厌恶、恐惧、惊讶或喜悦，以及中性类别。
 - 我们利用 one-hot 向量来表示与每条推文相对应的情感类别。

实验、结果和分析

下面介绍了对我们为所提出的方法引入的数据集和现有的最先进基线进行的实验、结果和分析的广泛讨论。

基线

以下基线方法与我们提出的方法进行了比较。

- $XLNet^{FT+LS+RF}$ 方法中，对预训练的 BERT（来自 Transformers 的双向编码器表示）模型进行微调 (FT) 以执行情感分析。为了减少过度拟合，作者结合了标签平滑 (LS) 和基于规则的功能 (RF)，例如否定处理和情绪转变检测。该模型用于表情符号、情感和情绪分析任务。
- 多语言 BERT (mBERT)：作者利用一种名为 mBERT 的基于转换器的语言模型来学习多种语言中单词的上下文嵌入。mBERT 在大量单语言和多语言文本数据上进行了预训练，并在 SentiMix 代码混合数据集上进行了微调，用于情感检测和情感识别。
- $XLNet^{MTL}$ 作者使用了 XLM-R，这是一种基于 Transformer 架构的跨语言语言模型，该模型在包括代码混合文本的更大数据集上进行了预训练。XLM-R 可以对多种语言的文本进行编码和解码，并在各种 NLP 任务上取得了最先进的结果，包括情感分析和情感识别。他们在 SentiMix 代码混合数据集上对 XLM-R 进行了微调，以进行情感检测和情感识别。
- $TL-XLMR^{LS}$ 检测 SentiMix 代码混合数据集中的情绪并识别情绪，作者采用了基于 Transformer 架构的端到端多任务框架。他们使用特定于任务的数据对预训练的跨语言嵌入模型 XLMRoBERTa (XLMR) 进行了微调，以通过迁移学习提高模型效率。
- $TL-XLMR^{LS}$ 在这个消融实验中，作者用 mBERT 替换了 XLMR 模块，以研究句子编码器在 TL-XLMR 中的意义。该模型在 SentiMix 代码混合数据集上进行了微调，以执行情感检测和情感识别。

实施细节

我们建议的模型使用 PyTorch 付诸实践，PyTorch 是一个很受欢迎的 Python 深度学习工具包。我们使用 F1 分数 (F1) 作为情绪和情绪预测的评估指标，对于表情符号，我们使用 Jaccord 指数 (JI)、宏观 F1 分数。我们利用 Adam 优化器并进行 200 个 epoch 的网格搜索来改进模型。我们使用具有两层的 Transformer 编码器，我们根据经验发现嵌入大小为 300（检查了 100、150、200 和 300）。Dropout 率设置为 0.5，而学习率设置为 0.05。使用经验技术发现自动编码器的维度为 2048。判别器，由两个全连接层 (ReLU 层) 组成。学习率设置为 $1e-3$ ，权重衰减为 $1e-4$ ，动量为 0.3。通过将 F1 和准确度分数与不同基线进行对比，可以评估我们策略的有效性。在 CM-RFT 中，内核是使用全连接层根据输入动态计算的。内核大小为 $[3, 5, 7, 31 \times 3]$ ，每个模块有 4 个头 (Transformer 基本模型中头数量的一半)。

评估指标

对于表情符号检测任务，我们考虑杰卡德指数（JI）和汉明损失（HL）指标来评估我们提出的系统的性能。此外，我们还报告了微平均 F1score 和准确度值（如表 8 所示）。JI、HL 和微平均 F1 是评估多标签分类任务的常用选择。对于情绪和情绪检测任务（如表 9 和表 10 所示），我们报告了我们提出的模型的宏观平均 F1 分数和准确度值。

- 微平均F1分数：对于多标签分类任务，微平均F1分数是一种常用的指标，通过计算真阳性（TP）、假阴性（FN）和假阳性（FP）来全局计算F1分数）跨越所有标签。微平均F1分数的公式为：
$$F1 = \frac{\sum_{i=1}^n \frac{2 * TP_i}{2 * TP_i + FP_i + FN_i}}{\sum_{i=1}^n 1}$$
- 宏观平均 F1 分数：宏观平均 F1 分数是多标签分类任务的另一个常用指标。它计算每个标签的 F1 分数，然后取这些 F1 分数的平均值。宏观平均F1分数的公式为：
$$F1 = \frac{\sum_{i=1}^n \frac{2 * TP_i}{2 * TP_i + FP_i + FN_i}}{n}$$
- 准确率：准确率是衡量正确分类的标签占标签总数比例的指标。准确度的公式为：
$$A = \frac{\sum_{i=1}^n TP_i}{\sum_{i=1}^n TP_i + \sum_{i=1}^n FP_i}$$
- 汉明损失：汉明损失衡量错误分类的标签占标签总数的比例。汉明损失的公式为：
$$HL = \frac{\sum_{i=1}^n m_i}{n}$$
 其中 n 是实例数，m 是

| 任务 | | 嵌入 | | | CM-T | | | | CM-FT | | | | CM-RFT | | | |
|-----|--------|----|------|--------|--------|-------|------|------|--------|-------|-------|------|--------|-------|-------|------|
| | | CE | Elmo | TF-IDF | F1 (%) | 率 (%) | HL | JI | F1 (%) | 率 (%) | HL | JI | F1 (%) | 率 (%) | HL | JI |
| UTL | mE | √ | √ | - | 0.56 | 0.58 | 0.19 | 0.43 | 0.57 | 0.60 | 0.19 | 0.45 | 0.59 | 0.62 | 0.15 | 0.52 |
| | mE | √ | √ | √ | 0.57 | 0.59 | 0.18 | 0.44 | 0.60 | 0.62 | 0.18 | 0.47 | 0.62 | 0.65 | 0.14 | 0.54 |
| | mE | √ | √ | √ | 0.59 | 0.62 | 0.18 | 0.46 | 0.63 | 0.65 | 0.17 | 0.49 | 0.64 | 0.67 | 0.13 | 0.56 |
| DTL | S+mE | √ | √ | - | 0.58 | 0.61 | 0.16 | 0.53 | 0.60 | 0.63 | 0.13 | 0.55 | 0.63 | 0.65 | 0.11 | 0.58 |
| | S+mE | √ | √ | √ | 0.61 | 0.63 | 0.15 | 0.52 | 0.64 | 0.67 | 0.12 | 0.56 | 0.65 | 0.69 | 0.09 | 0.60 |
| | S+mE | √ | √ | √ | 0.62 | 0.64 | 0.14 | 0.54 | 0.65 | 0.68 | 0.11 | 0.58 | 0.68 | 0.71 | 0.07 | 0.61 |
| | 米+E | √ | √ | - | 0.59 | 0.60 | 0.17 | 0.47 | 0.60 | 0.63 | 0.15 | 0.50 | 0.60 | 0.64 | 0.13 | 0.55 |
| | 米+E | √ | √ | √ | 0.62 | 0.63 | 0.16 | 0.49 | 0.62 | 0.64 | 0.14 | 0.52 | 0.64 | 0.66 | 0.11 | 0.57 |
| TTL | 米+E | √ | √ | √ | 0.61 | 0.64 | 0.14 | 0.51 | 0.64 | 0.67 | 0.13 | 0.54 | 0.67 | 0.69 | 0.09 | 0.59 |
| | S+E+mE | √ | √ | - | 0.60 | 0.65 | 0.12 | 0.56 | 0.63 | 0.67 | 0.064 | 0.62 | 0.67 | 0.74 | 0.059 | 0.64 |
| | S+E+mE | √ | √ | √ | 0.61 | 0.65 | 0.10 | 0.59 | 0.65 | 0.68 | 0.061 | 0.64 | 0.69 | 0.73 | 0.057 | 0.67 |
| | S+E+mE | √ | √ | √ | 0.63 | 0.66 | 0.09 | 0.60 | 0.67 | 0.70 | 0.056 | 0.66 | 0.73 | 0.75 | 0.054 | 0.69 |
| | S+E+mE | √ | √ | √ | 0.63 | 0.66 | 0.09 | 0.60 | 0.67 | 0.70 | 0.056 | 0.66 | 0.73 | 0.75 | 0.054 | 0.69 |

表 8.我们提出的多标签表情符号分类 CM-RFT 框架的结果。最高分数以粗体显示。

：多标签表情符号分类，S：情感分类，mE：情绪分类，

类，CE：字符嵌入，A：准确性，

HL：汉明损失，

JI：杰卡德指数。

| 任务 | | 嵌入 | | | CM-T | | CM-FT | | CM-RFT | |
|-----|--------|----|------|--------|--------|-------|--------|-------|--------|-------|
| | | CE | Elmo | TF-IDF | F1 (%) | 率 (%) | F1 (%) | 率 (%) | F1 (%) | 率 (%) |
| UTL | S | √ | √ | - | 64.21 | 64.93 | 65.89 | 67.11 | 69.49 | 71.55 |
| | S | √ | √ | √ | 66.69 | 68.17 | 68.32 | 68.79 | 71.88 | 73.34 |
| | S | √ | √ | √ | 68.24 | 69.34 | 70.11 | 71.78 | 72.65 | 75.19 |
| | S+mE | √ | √ | - | 67.26 | 71.82 | 69.43 | 72.91 | 74.69 | 76.71 |
| DTL | S+mE | √ | √ | √ | 70.31 | 71.81 | 71.72 | 73.48 | 76.41 | 77.35 |
| | S+mE | √ | √ | √ | 71.42 | 74.34 | 73.34 | 75.51 | 78.22 | 79.21 |
| | S + E | √ | √ | - | 66.13 | 68.79 | 68.81 | 70.39 | 72.21 | 74.33 |
| | S + E | √ | √ | √ | 67.71 | 70.19 | 69.29 | 71.32 | 73.98 | 75.87 |
| | S + E | √ | √ | √ | 69.83 | 71.98 | 71.31 | 73.10 | 74.64 | 77.31 |
| TTL | S+E+mE | √ | √ | - | 71.98 | 72.79 | 73.31 | 74.53 | 78.54 | 79.27 |
| | S+E+mE | √ | √ | √ | 72.51 | 75.35 | 74.89 | 77.49 | 81.11 | 81.93 |
| | S+E+mE | √ | √ | √ | 73.31 | 78.43 | 75.42 | 79.26 | 82.35 | 83.14 |
| | S+E+mE | √ | √ | √ | 73.31 | 78.43 | 75.42 | 79.26 | 82.35 | 83.14 |

表 9.我们提出的用于情感分类的 CM-RFT 框架的结果。最先进的情绪结果为 0.75，如中所述。最高分数以粗体显示。

：情感分类，

：多标签表情符号分类，CE：字符嵌入，A：准确性，HL：汉明损失，JI：杰卡德指数。

| 任务 | | 嵌入 | | | CM-T | | CM-FT | | CM-RFT | |
|-----|--------|----|------|--------|--------|-------|--------|-------|--------|-------|
| | | CE | Elmo | TF-IDF | F1 (%) | 率 (%) | F1 (%) | 率 (%) | F1 (%) | 率 (%) |
| UTL | E | √ | √ | - | 46.21 | 48.54 | 47.54 | 49.22 | 48.78 | 52.77 |
| | E | √ | √ | √ | 48.53 | 50.11 | 50.86 | 52.67 | 52.31 | 54.54 |
| | E | √ | √ | √ | 50.32 | 52.79 | 52.67 | 54.72 | 54.21 | 56.79 |
| | S + E | √ | √ | - | 52.84 | 53.26 | 52.91 | 54.19 | 54.82 | 57.73 |
| DTL | S + E | √ | √ | √ | 52.98 | 55.03 | 54.41 | 56.91 | 55.41 | 58.91 |
| | S + E | √ | √ | √ | 53.92 | 56.31 | 55.23 | 58.92 | 57.33 | 61.22 |
| | 米+E | √ | √ | - | 48.42 | 51.12 | 50.32 | 52.81 | 52.83 | 55.95 |
| | 米+E | √ | √ | √ | 50.42 | 53.80 | 52.72 | 54.30 | 53.25 | 56.15 |
| TTL | 米+E | √ | √ | √ | 51.21 | 54.79 | 54.11 | 56.90 | 55.76 | 58.63 |
| | S+E+mE | √ | √ | - | 53.88 | 57.26 | 56.19 | 58.37 | 57.87 | 59.54 |
| | S+E+mE | √ | √ | √ | 55.21 | 57.82 | 57.51 | 59.71 | 59.33 | 61.91 |
| | S+E+mE | √ | √ | √ | 56.79 | 59.18 | 58.71 | 61.63 | 60.53 | 63.73 |

表 10.我们提出的情绪分类 CM-RFT 框架的结果。最高分数以粗体显示。
S：情感分类，E：情绪分类，mE：多标签表情符号分类，CE：字符嵌入，A：准确性。

标签数量， \mathbf{Y} 是真实的标签向量，例 i ， $\hat{\mathbf{Y}}$ 是实例 i 的预测标签向量，并且 xor 是逻辑异或运算符。

- 杰卡德指数：杰卡德指数通过计算两个集合的交集大小与并集大小的比率来衡量两个集合之间的相似度，用于衡量多标签分类中预测标签集与真实标签集之间的相似度。杰卡德指数的公式为：
$$JI = \frac{|\mathbf{Y} \cap \hat{\mathbf{Y}}|}{|\mathbf{Y} \cup \hat{\mathbf{Y}}|}$$
其中 \mathbf{Y} 是实例 i 的真实标签集。Jaccard 相似度计算为预测标签集和真实标签集交集的大小除以它们并集的大小。所得分数范围从 0 到 1，其中 1 表示预测标签集与真实标签集之间的完美相似性。

结果

表 8、9 和 10 显示了 CM-T、CM-FT 和 CM-RFT 模型在 UTL、DTL 和 TTL 设置中的表情符号、情绪和情绪任务的性能。这些设置研究了与单任务学习相比，多任务学习在提高整体系统性能方面的有效性。

多标签表情符号分类（mE）
表 8 中报告的结果是在三种不同设置（单任务学习、双任务学习和三任务学习）上训练的三种不同模型（CM-T、CM-FT、CM-RFT）的性能指标表情符号检测的任务。
在单任务学习设置中，每个任务都是单独解决的，CM-RFT 模型的性能随着更多特征的添加而提高。具体来说，当我们从仅使用字符嵌入到字符嵌入 + Elmo 嵌入 + TF-IDF 时，性能会有所提高。F1 分数从 0.59 增加到 0.64，准确率从 0.62 增加到 0.67，汉明损失从 0.15 减少到 0.13，Jaccard 指数从 0.52 增加到 0.56。这些结果表明，使用多个特征可以提高表情符号检测任务的性能。
在双任务学习设置中，表情符号任务与情绪/情绪任务联合学习，与单任务学习设置相比，CM-RFT 模型的性能进一步提高。当模型接受字符嵌入 + Elmo 嵌入 + TF-IDF 特征训练时，改进更加明显。F1 分数从 0.64 增加到 0.68，准确率从 0.67 增加到 0.71，汉明损失从 0.13 减少到 0.07，Jaccard 指数从 0.56 增加到 0.61。这些结果表明，在多个任务上训练模型可以进一步提高表情符号检测任务的性能。
在联合学习情感、情绪和表情符号检测任务的三任务学习设置中，与双任务学习设置相比，CM-RFT 模型的性能进一步提高。F1 分数从 0.68 增加到 0.73，准确率从 0.71 增加到 0.75，汉明损失从 0.07 减少到 0.054，Jaccard 指数从 0.61 增加到 0.69。这些结果表明，多个任务的联合学习可以显着提高表情符号检测任务的性能。
总体而言，结果表明，可以通过使用多个特征并在多个任务上训练模型来提高表情符号检测任务的性能。此外，结果表明情绪和情感对表情符号检测任务的性能有显着影响，因为这些任务的联合学习可以显着提高性能。

S)

情感分类 (

表9列出了情感和表情符号任务联合学习的情感分类任务结果。在单任务设置中，每个任务都是独立执行的，CM-RFT 模型在情感任务中实现了最高性能，F1 得分为 72.65，准确度为 75.19。这表明，与仅使用字符嵌入特征的模型相比，包含额外特征（例如 Elmo 嵌入和 TF-IDF 特征）可以增强所有模型的情感检测性能。

在双任务设置中，当联合学习情感和表情符号任务时，情感检测任务的 F1 分数和准确度分数从单任务设置中的 72.65 和 75.19 分别提高到 78.22 和 79.21。字符嵌入、Elmo 嵌入和 TF-IDF 特征。类似地，当联合学习情感和表情符号任务时，当使用字符嵌入、Elmo 嵌入时，情感检测任务的 F1 分数和准确度分数分别从单任务设置中的 72.65 和 75.19 提高到 74.64 和 77.31，以及 TF-IDF 特征。

在联合解决情感、情绪和表情符号检测任务的三任务设置中，CM-RFT 模型在情感任务中实现了最佳性能，F1 得分为 82.35，准确度为 83.14，其次是 CM-FT模型的 F1 得分为 75.42，准确度为 79.26。这再次证实，当多任务学习与其他任务联合学习时，有助于提高情感检测性能。

研究结果表明，将情感和表情符号检测任务集成到情感分类任务中可以提高模型的性能。三任务学习设置展示了情感任务的最高性能，这意味着合并这些额外任务可以提高模型对文本中表达的情感的理解。性能的增强可能是由于情绪和表情符号提供了额外的上下文信息，特别是在情绪复杂或讽刺的情况下。因此，结合情感和表情符号检测任务可能是增强情感分类模型性能的有用技术。此外，结合其他功能，例如 Elmo 嵌入和 TF-IDF 功能，也可以提高情感检测性能。

E):

情绪分类 (

根据表 10 中的结果，我们可以观察到，随着我们从单任务学习过渡到双任务学习，最终过渡到三任务学习，情感任务的表现会有所提高。在单任务设置中，CM-RFT 模型在所有三种特征组合上都优于 CM-T 和 CM-FT 模型，这表明合并情感和表情符号信息可以提高情绪检测任务的性能。在带有表情符号的双任务设置中，所有模型的性能都远低于单任务设置。然而，随着更多特征的加入，性能得到提高，并且 CM-RFT 模型在所有三个特征上都取得了最佳结果。这表明利用各种特征类型可以有利于表情符号和情绪检测的联合学习，并且三任务设置可以提供进一步的改进。在带有情感的双任务设置中，性能优于表情符号。Elmo 嵌入和 TF-IDF 特征的添加带来了一致的性能改进，CM-RFT 模型再次取得了最佳结果。这意味着情感和情绪检测的联合学习也可以从多种特征类型的使用中受益。

结果表明，情感和表情符号信息的存在似乎可以增强情感任务的表现。情感任务的最佳性能是在三任务学习设置中获得的，其中涉及联合学习情感、情感和表情符号检测任务。性能的提高可归因于情绪和表情符号提供了额外的上下文信息，有助于更好地消除情绪歧义。

结果还表明，多任务学习比单任务学习更有效，特别是当任务相关时，例如情感、情感和表情符号检测。随着我们从单任务学习到双任务学习，最后发展到三任务学习，情感任务的表现持续提高。这表明相关任务的联合学习可以更好地利用可用信息并提高系统的整体性能。

与现有作品的比较

表 11 中的结果表明，本研究中提出的 CM-RFT 模型在情感和表情符号检测任务方面都比最先进的模型表现更好。在单任务场景中，mBERT 在表情符号检测任务中取得了最高的 63.77% 准确率和 61.54% 的 F1 分数。然而，在多任务设置中，所提出的 CM-RFT 模型超越了所有其他模型，达到了 75.81% 的准确率和 73.25% 的 F1 分数。这表明所提出的模型有效地利用多任务学习来提高两项任务的性能。此外，该模型在无监督情绪检测任务中也显示出了可喜的结果，F1 得分为 60.53%，准确率为 63.73%。这表明，即使没有标记数据，所提出的模型中使用的零样本方法也能有效地从文本中检测情感。

当关注表情符号预测任务时，所提出的 CM-RFT 模型显着优于单任务和多任务模型。该模型的准确率达到了 75.81%，比表现最佳的单任务模型（mBERT）的准确率高出约 12%，比表现最佳的多任务模型（TL-）的准确率高出约 9%。XLMR

）。此外，该模型的 F1 分数为 73.25%，比表现最佳的单任务模型^{LS1}（mBERT）的 F1 分数高出约 12%，比表现最佳的多任务模型的 F1 分数高出约 8% (TL-XLMR^{LS1})^{LS1}。

我们对提出的模型进行了额外的实验，以将其与前面讨论的单任务和多任务基线进行公平比较。由于没有一个基线模型解决无监督分类问题，因此它们无法为情感任务生成分数，这与我们提出的 CM-RFT 模型不同，该模型解决了监督环境中的情感和多标签表情符号检测以及无监督环境中的情感检测，使用

| 型号 | 表情符号 | | 情绪 | | 情感 | |
|---|--------|-------|--------|-------|--------|-------|
| | F1 (%) | 率 (%) | F1 (%) | 率 (%) | F1 (%) | 率 (%) |
| 单任务基线 | | | | | | |
| XLMR | 58.18 | 61.21 | 67.42 | 70.38 | - | - |
| XLMR | 59.89 | 61.04 | 68.55 | 71.43 | - | - |
| mBERT | 61.54 | 63.77 | 70.32 | 73.21 | - | - |
| CM-RFT ^{STL*} _[-情绪] (建议的) | 64.14 | 67.30 | 74.64 | 77.31 | 55.76 | 58.63 |
| 多任务基线 | | | | | | |
| XLMR | 63.52 | 64.88 | 72.58 | 75.44 | - | - |
| TL-XLMR | 65.30 | 66.83 | 78.21 | 80.55 | - | - |
| TL-mBERT | 64.68 | 65.77 | 76.59 | 78.23 | - | - |
| CM-RFT ^{MTL*} _[-情绪] (建议的) | 68.57 | 71.68 | 78.22 | 79.21 | - | - |
| 多任务-Zeroshot-Emotion-BenchMark | | | | | | |
| CM-RFT* (建议) | 73.25 | 75.81 | 83.14 | 82.35 | 60.53 | 63.73 |

表 11.我们提出的模型与其他基线之间的比较。最高分数以粗体显示。标有 * 的结果是建议任务的基准结果。

零射击方法。因此，我们训练了两个版本的 CM-RFT 模型：一个在单任务设置中 (CM-RFT^{STL}_[-情绪]) 对于所有任务以及多任务设置中的另一个 (CM-RFT^{MTL}_[-情绪]) 没有情感任务。结果如表 11 所示。CM-RFT 性能观察到 STL-CM-RFT 在表情符号和情绪任务的准确性和 F1 分数方面优于所有这些模型。例如，CM-RFT 的准确率为 67.30%，而 mBERT 单 emoji 任务的准确率为 63.77%，而单任务模型的最高 F1 分数为 mBERT 的 70.32%。这些结果表明，纳入无监督情感任务确实有助于模型在监督任务上表现更好。

CM-RFT 性能比较在表情符号和情绪任务的准确性和 F1 分数方面优于所有这些模型。例如，CM-RFT 的准确率为 67.30%，而 mBERT 单 emoji 任务的准确率为 63.77%，而单任务模型的最高 F1 分数为 mBERT 的 70.32%。这些结果表明，纳入无监督情感任务确实有助于模型在监督任务上表现更好。我们的模型在情感识别任务上的性能。我们的模型在情感识别方面的 F1 得分为 60.53，与 Llama 模型的 F1 得分非常接近，后者的 F1 得分为 61.11。这些结果强调了我们提出的方法在处理情感识别任务方面的有效性，表明其在自然语言处理中的实际应用潜力。综上所述，我们提出的 CM-RFT 模型在情感和表情符号检测任务上都优于当前最先进的模型。我们的结果表明，利用多任务学习和利用零样本方法进行无监督情绪检测可以显着提高任务性能。对于表情符号预测任务，我们提出的模型比表现最好的单任务和多任务模型取得了显着的改进，证明了我们方法的有效性。

与各种数据集的当前模型进行比较

为了评估我们模型的有效性，我们与几篇论文及其相应模型进行了比较。

- 比较研究 1: “代码混合罗马乌尔都语 - 英文文本中的情绪检测”。模型：我们将我们的模型与 BERT 和 XLM-RoBERTa 进行了比较。使用的数据集：我们使用了 Code-Mixed Roman Urdu - English Text 数据集。结果如表 12 所示，表明我们的模型优于 BERT 和 XLM-RoBERTa，F1 得分为 0.69，证明了其在检测代码混合文本中的情绪方面的有效性。
- 比较研究 2: “代码混合语言的自注意力混合表情符号预测模型” 模型：我们将我们的模型与 BARF 进行了比较。使用的数据集：我们使用了印度英语表情符号预测 (HEP) 数据集。结果如表 13 所示，表明与 BARF 相比，我们的模型获得了更高的 F1 分数 0.64，证明了其在预测代码混合语言中表情符号方面的优越性能。
- 比较研究 3: “代码混合的印度英语数据中的情感检测和情感识别的多任务处理” 模型：我们将我们的模型与 TL-XLMR 进行了比较^{LS}。使用的数据集：我们使用 SemEval-2020 任务 9 数据集。表 14 显示了结果，表明我们的模型在两种情绪上都获得了更高的 F1 分数

| 模型 | F1分数(情感) |
|--------------|----------|
| BERT | 0.28 |
| XLM-罗伯塔 0.60 | |
| CM-RFT | 0.69 |

表 12. 代码混合罗马乌尔都语-英语文本数据集的比较。

| 模型 | F1分数(表情符号) |
|---------|------------|
| 呕吐 0.59 | |
| CM-RFT | 0.64 |

表 13. 印度英语表情符号预测 (HEP) 数据集的比较。

与表明其在代码混合检测(70.31)和情绪分析(70.31)数据中情感和情绪识别的多任务处理中的有效性。

MTL
LS

消融研究

对所提出的 CM-RFT 框架组件进行的消融实验
表 15 显示了四个消融实验的结果，旨在评估不同组件在所提出的 CM-RFT 框架中的贡献。检查的四个组件是 GLU 模块、自动编码器和 ANP 模块、自注意力机制以及 GLU、自注意力、ANP 和 AE 模块的集体组合。

结果表明，每个组件都有助于 CM-RFT 框架的整体性能。删除任何这些组件都会导致所有三个任务的 F1 分数显着下降，特别是当所有四个模块都被删除时（第 4 行）。这表明所提出的框架设计良好，每个模块对其成功都发挥着关键作用。具体来说，GLU 模块似乎是框架的关键部分（第 1 行）。删除该组件会导致所有三个任务的性能显着下降，凸显了文本编码器中非线性变换的重要性。同样，删除自动编码器和 ANP 模块会导致性能下降（第 2 行），这表明这些无监督预训练方法在学习有用的特征表示方面的重要性。此外，在融合 GLU 和 Trans Encoder 模块的输出特征方面，自注意力机制似乎比线性串联更有效（第 3 行）。这一结果证实了自注意力在捕获远程依赖性和对输入标记之间的交互进行建模方面的卓越性能。最后，GLU、SA、ANP 和 AE 模块的集体组合是一种高效的特征学习机制（第 4 行），但它也导致更高的计算成本。结果表明，尽管以模型容量和表达能力降低为代价，但仍然可以通过更简单的线性串联机制获得不错的性能。

| 模型 | F1 分数（情感） | F1 分数（情绪） |
|--------------------------|-----------|-----------|
| TL-XLMR ^{MTLLS} | 71.61 | 64.47 |
| CM-RFT | 76.22 | 70.31 |

表 14. SemEval-2020 Task 9 数据集的比较。

| 设置 | 表情符号 F1 (%) | 情感 F1 (%) | 情绪 F1 (%) |
|--------------------------------|----------------|--------------|--------------|
| [CM-RFT] | 69.97 (3.79) | 57.39 (3.41) | 78.54 (3.81) |
| [CM-RFT]-(AE+ANP) | 70.97 (2.79) | 58.15 (2.38) | 79.34 (3.01) |
| [CM-RFT]-SA | 71.78 (1.98) | 58.22 (2.31) | 79.69 (2.66) |
| [CM-RFT]-(GLU + SA + AE + ANP) | 66.97 (6.79) | 53.52 (7.01) | 75.37 (6.98) |
| [CM-RFT] | 73.76 | 60.53 | 82.35 |

表 15. 消融实验结果。分数下降百分比显示在括号中。GLU：门控线性单元，SA：自注意力融合，AE：自动编码器，ANP：形容词-名词对。

总之，消融实验证明了所提出的 CM-RFT 框架中每个模块对于多标签表情符号预测的重要性。研究结果可以指导未来模型的设计，并揭示有助于其成功的潜在机制。

损失函数对于模型训练的重要性

表 16 显示了四个消融实验的结果，其中每个实验都与包含表情符号、情感和情绪任务的所有三个损失函数（L、L 和 L）的建议 CM-RFT 进行比较。

当删除任何损失函数时，在每个消融实验中，所有三个任务的 F1 分数持续下降。当所有三个损失函数被删除时，性能下降幅度最大，这表明每个损失函数对模型的性能都起着重要作用。具体来说，删除

L 与仅删除其中一个损失函数相比，Lloss 函数对模型性能的负面影响最大。这表明这些损失函数对模型捕获对抗训练和输入数据重建的相关特征的能力有很大贡献。

就单个损失函数的贡献而言，与对齐损失（L）和重建损失（L）相比，对抗性损失（L）似乎对性能有稍大的影响，特别是对于表情符号和情绪检测任务。这表明对抗性损失在模型区分这些任务的不同类别的能力中起着重要作用。另一方面，对齐损失和重建损失对于情感检测似乎更重要。

总的来说，这些结果证明了所提出的损失函数对于有效训练多任务表情符号、情感和情绪检测系统的重要性。这些发现可用于指导未来为多任务学习模型开发更有效的训练策略。例如，合并额外的损失函数或修改现有损失函数的权重可以提高模型的性能。此外，这些结果表明，不同损失函数的重要性可能会根据所执行的具体任务和所使用的数据而有所不同，这凸显了在多任务学习模型设计中仔细分析和选择损失函数的重要性。

定性分析

在本节中，我们对所提出的多任务框架进行定性分析，该框架考虑了表情符号、情绪和情感之间的关系，正如我们之前提到的。为了说明这些任务相互之间的影响，我们从 SENTIMOJI 数据集中选择了几个示例，并将它们呈现在表 17 中。

- 观察 1: 在第一句话中，模型正确预测了心形表情符号、积极情绪和喜悦作为情感。该模型似乎分别从“太好了”和“不知道”这两个词中汲取了积极情绪和喜悦，并预测了心形表情符号以匹配积极情绪。此外，“bhai”（兄弟）一词可能暗示着友好或深情的语气，从而导致心形表情符号的识别。最后，训练数据中单词“joy”或类似单词的存在可能有助于模型准确识别情绪。
- 观察2: 在第二句话中，模型正确预测了负面情绪，但预测的表情符号是错误的。该模型预测的是一张撅嘴的脸，而不是一张愤怒的脸，这可能是因为撅嘴的表情符号也可以表示不满或烦恼，这可能与骄傲有关。此外，这种情绪被错误地归类为厌恶而不是愤怒，这可能是因为强烈的负面情绪以及使用“失败”和“不能这样做”等词语。
- 观察 3: 在第三句话中，模型正确预测了张开嘴、呕吐的表情符号，表示厌恶，以及负面情绪。该句子包含“失踪”等词语，暗示着一种消极情绪，而张开嘴、呕吐表情符号和厌恶情绪的使用可能与句子中表达的厌恶情绪有关。
- 观察 4: 在第一个多标签句子中，模型正确预测了负面情绪和喜悦作为情绪，但仅部分预测了表情符号。句子中使用“hardik subhkamnaye”和“Congratulations Sir ji”表示积极情绪，使用“Dobara pm banee”表示成就感，这可以解释心形和闪光表情符号的使用。笑脸表情符号的错误分类可能是由于缺乏上下文信息或训练数据不足造成的。

| 基本型号 | L | L | L | F1 (%) | 情感 F1 (%) | 情绪 F1 (%) |
|------|---|---|---|---------------|---------------|--------------|
| ✓ | - | - | - | 69.60 (-4.16) | 56.09 (-4.44) | 78.08(-4.27) |
| ✓ | ✓ | - | ✓ | 72.25 (-1.51) | 58.96 (-1.57) | 80.52(-1.83) |
| ✓ | - | ✓ | ✓ | 71.73 (-2.03) | 58.74 (-1.79) | 80.14(-2.21) |
| ✓ | - | - | ✓ | 70.77 (-2.99) | 57.22 (-3.31) | 78.23(-4.12) |
| ✓ | ✓ | ✓ | ✓ | 73.76 | 60.53 | 82.35 |

表 16. 不同损失函数对模型的影响。在这里，基本模型定义了我们的基本结构，其中语义特征通过零样本损失函数直接组合。最高分数以粗体显示。重要值以粗体显示。

| 单标签表情符号预测 | | | | | | | | |
|---|---|---------|-----|-----|---------|-----|-----|--|
| 代码混合推文 | 英文翻译 | 实际的 | | | 预测 | | | |
| | | 表情符号 | 情感 | 情绪 | 表情符号 | 情感 | 情绪 | |
| 我不知道 tu itna acha kaise hai bhai aaj bhi | 我不知道你今天怎么太好了 | ❤️ | Joy | 积极的 | ❤️ | Joy | 积极的 | |
| tere ghamand k karan hi aaj congress k ye halat hai ... failure hai tu Bhai .. Tujhse na ho payega | 因为你们的骄傲，这就是国会今天的状况。 ..你失败了..你不能这样做 | 😡 | 愤怒 | 消极的 | 😞 | 厌恶 | 消极的 | |
| 国会 ki sarker mai 汽缸 he gayab ho gaya tha | 在国会期间政府。气缸失踪 | 😡 | 厌恶 | 消极的 | 😞 | 厌恶 | 消极的 | |
| 多标签表情符号 | | | | | | | | |
| 代码混合推文 | 英文翻译 | 实际的 | | | 预测 | | | |
| | | 表情符号 | 情感 | 情绪 | 表情符号 | 情感 | 情绪 | |
| meta RT @narendramodi 恭喜 ji Dobara 先生 pm banee ki Hardik subhkamnaye aapko | meta RT @narendramodi 恭喜先生第二次成为总理。 | 😊, ✨, 🍷 | Joy | 积极的 | ❤️, ✨ | Joy | 积极的 | |
| 元 RT @Chaosthestic 其次, maa baap ki izzat b koi cheez hoti hai. Har cheez ko 支持卡纳乐队卡罗. 仅仅因为她是女性并不意味着我是女人 | meta RT @Chaosthestic 其次, 父亲和母亲的尊重是重要的。不支持一切。仅仅因为她是女性并不意味着她可以说什么。 | 😡, 😞, 🍷 | 愤怒 | 消极的 | 😞, 😞, 🍷 | 厌恶 | 消极的 | |
| 我真的很想见到他, 用我的手表达我的爱.. saale ne jeena muskil kar rkha hai.. | 我真的很想见到他, 用我的双手表达我的爱..他让我的生活变得悲惨.. | 😞, 😡, 🍷 | Sad | 消极的 | 😡, 😞 | 愤怒 | 消极的 | |

表 17. 拟议框架对表情符号、情感和情绪的预测如下所示。

- 观察 5: 在第二个多标签句子中, 模型正确预测了负面情绪, 但将情绪错误地分类为厌恶而不是愤怒。对于表情符号, 模型预测了撅嘴脸、哭泣脸和失望脸, 但原始注释有撅嘴脸、愤怒脸和张开嘴、呕吐的脸。这可能是因为模型识别出了负面情绪以及“尊重”、“任何事情”和“女人”等词语的使用, 这可能导致预测出撅嘴的表情符号, 而预测出哭泣的表情和失望的表情。面部表情符号可能与负面情绪有关。
- 观察 6: 在第三个多标签句子中, 模型正确地将情绪识别为负面, 但错误地将情绪预测为愤怒而不是悲伤。该模型还部分预测了表情符号, 这可能是由于句子中存在多种情绪。为了改进预测, 可以对模型进行更多包含相似短语和单词的数据的训练, 以更好地区分不同的负面情绪和表情符号。

误差分析

对错误预测的分析揭示了几种常见的错误模式, 总结如下:

- 表情符号解释中的歧义: 该模型经常会遇到根据上下文有多种解释的表情符号。例如, 表情符号既可以代表欢笑, 也可以代表喜悦的泪水, 从而导致错误分类。
- 否定和讽刺: 文本中的否定和讽刺可能会导致模型产生误解, 尤其是在情感分析中。例如, 短语“不错”可能会被模型解释为积极的, 从而导致错误分类。
- 缺乏上下文: 模型有时无法捕获句子的上下文, 导致情感和情绪分类错误。例如, 短句或上下文模糊的句子可能会被错误分类。
- 数据不平衡: 类别分布的不平衡可能会导致模型预测出现偏差, 特别是对于少数类别。这在情感分类中尤其明显, 其中某些类别的示例少于其他类别。
- 词汇表外的单词: 文本中出现词汇表外的单词可能会导致错误, 尤其是当模型无法捕获其语义时。这在表情符号和情感分析任务中更为常见。

这些错误模式凸显了所提出的 CM-RFT 模型在跨不同任务理解和解释文本时面临的挑战。应对这些挑战需要进一步研究更稳健的建模技术、更好地处理上下文和歧义以及减轻数据偏差。

情绪和情感任务与表情符号预测任务的联合学习可能有利于表情符号任务的性能。这是因为情绪和情绪可以为模型提供额外的上下文来预测适当的表情符号。例如, 在第一个正确的预测样本中, 模型能够正确预测心形表情符号, 这可能受到该句子预测的积极情绪和快乐情绪的影响。同样, 在第二个错误预测样本中, 模型正确

预测了负面情绪，但对情绪和表情符号进行了错误分类，这表明它可能没有完全捕捉到文本的细微差别。单标签表情符号在多标签表情符号预测中可能存在风险，因为表情符号在不同的上下文中可能具有不同的含义，并且单个表情符号可能无法捕获文本的所有细微差别。例如，“撇嘴”表情符号可用于表达愤怒、失望或悲伤，如果没有额外的上下文，可能很难确定所传达的确切情绪。我们在错误的预测样本中观察到，模型正确预测了一些表情符号，但遗漏了一些表情符号。这比完全错误的预测要好，因为它表明模型对上下文有一定的理解，并且可以在一定程度上预测相关的表情符号。然而，该模型的性能仍有改进的空间。

为了改进模型的预测，我们可以考虑以下步骤：

- 增加训练数据：模型可能会受益于额外的训练数据来捕获语言和情感的各种细微差别。
- 合并上下文：模型可能会受益于合并句子的上下文，以更好地识别情感、表情符号和情感。
- 使用预先训练的语言模型：该模型可能会受益于使用可以捕获单词和短语的语义的预先训练的语言模型。
- 正则化模型：模型可能会受益于正则化技术，以防止过度拟合并提高泛化能力。
- 分析和纠正错误：分析模型的错误并纠正它们可能有助于随着时间的推移提高模型的性能。

与 ChatGPT-3.5 的比较

我们使用 ChatGPT(<https://chat.openai.com/>) 进行一项研究来证明我们提出的框架的有效性。我们注意到 CM-RFT 比 ChatGPT 具有压倒性的性能优势。ChatGPT 对 TASKS 任务的一些预测示例如下所示：

提示：阅读这些印度英语话语并找到合适的表情符号、情感和情绪：

- tere liye chand nhi la sakta 宝贝实际上 tu bhaad mein ja
- Tere ghamand k karan hi aaj congress k ye halat hai ... failure hai tu Bhai .. Tujhse na ho payega
- 国会 ki sarker mai 汽缸 he gayab ho gaya tha

人类注释者：
表情符号标签 🤨、😏、😬 情感标签：愤怒、愤怒、厌恶。情绪标签：负面、负面、负面
Propose_MODEL：表情符号标签：、、、情绪标签：愤怒、厌恶、厌恶。情绪标签：消极、消极、消极 ChatGPT：😬表情符号标签：、、、情绪标签：轻蔑、愤怒、困惑。



情绪标签：负面、负面、中性（根据上下文，可能会被解释为负面）

在我们的分析中，很明显我们的模型产生类似于 ChatGPT 的结果。虽然 ChatGPT 以其高性能而闻名，但我们的模型表现出了熟练程度，特别是在处理代码混合句子方面。

局限性和潜在偏见

虽然我们提出的 CM-RFT 模型在多个任务中表现出强大的性能，但仍存在一些限制和潜在的偏差需要解决：

1. 数据偏差：模型的性能很大程度上依赖于训练数据的质量和代表性。训练数据中存在的偏差（例如某些人口统计数据或主题的代表性不足）可能会导致模型的预测出现偏差。
2. 语言偏差：由于语言结构、文化差异和训练数据可用性的差异，不同语言的模型性能可能会有所不同。与那些在训练数据中没有得到充分体现的语言相比，它在训练数据中得到充分体现的语言上可能会表现得更好。
3. 上下文敏感性：模型的性能受到文本呈现上下文的影响。它可能会遇到上下文模糊或讽刺性的文本，从而导致误解。
4. 泛化：模型泛化到未见过的数据或领域的能力受到训练数据的多样性和代表性的限制。它可能在与训练数据类似的数据上表现良好，但在处理域外或对抗性示例时表现不佳。
5. 可解释性：所提出的 CM-RFT 模型的复杂架构可能会阻碍其可解释性，使得理解某些预测的方式和原因变得困难。这种缺乏可解释性

透明度和问责制很重要的现实世界应用中，模型的有效性可能会受到限制。

解决这些局限性和偏见需要仔细考虑模型设计、训练数据、评估指标和道德考虑。未来的研究应侧重于开发更强大、更公平的人工智能模型，这些模型能够处理不同的语言、文化和环境，同时确保透明度、可解释性和问责制。此外，应努力收集更加多样化和有代表性的培训数据，并制定考虑偏见和公平问题的评估指标。通过解决这些挑战，我们可以构建对于现实世界应用来说更可靠、更公平、更值得信赖的人工智能模型。

结论

总之，我们的研究提出了一项新颖的任务，即预测代码混合句子中的多个表情符号，同时识别用户相应的情绪和情绪。我们引入了零样本策略，将情感信息整合到训练过程中，并开发了一个端到端框架来同时识别情绪、情绪和相应的表情符号。我们还通过扩展代码混合推文的基准情感注释 SentiMix 数据集，以及多标签表情符号的手动注释，引入了 SENTIMOJI 数据集。我们提出的多任务框架可以通过提供对表情符号、情感和代码混合句子中情感之间关系的全面理解来帮助弥合情感和情感分析领域的研究空白。我们的工作为未来的研究奠定了基础，以改善自然语言处理任务，这些任务需要更深入地理解代码混合文本在语言、情感和表情符号之间复杂的相互依赖关系。所提出的方法可以扩展到其他语言和领域，并且可以集成到社交媒体平台中，以改进代码混合文本的表情符号预测和情感分析。

我们工作的一个限制是我们没有探索使用更先进的预训练技术，例如自我监督学习。通过将这些技术融入到我们的模型中，我们可以增强模型学习表示的方式，并有可能提高其性能。尽管我们取得了最先进的结果，但通过探索更复杂的架构和进一步微调超参数，仍有改进的空间。在未来的研究中，我们提出的方法的有效性可以在不同语言的其他代码混合数据集上进行测试，以评估其泛化能力。研究我们的模型向其他语言的可迁移性在资源匮乏的环境中的表现如何，这将是一件很有趣的事情。最后，我们的方法可以扩展到其他 NLP 任务，例如代码混合语言的命名实体识别、机器翻译和语音识别。

数据可用性

这项研究涉及公开可用的 SentiMix 数据集 (<https://zenodo.org/record/3974927#.ZEva4pFBxH6>)，作者没有违反任何版权规范。代码和数据将在 <https://www.iitp.ac.in/~ai-nlp-ml/resources.html#SENTIMOJI> 提供

收稿日期：2023 年 4 月 30 日；接受日期：2024 年 4 月 4 日
Published online: 28 May 2024

参考

1. 巴比里, F.等人。Semeval 2018 任务 2: 多语言表情符号预测。第十二届语义评估国际研讨会论文集, 24-33 (2018)。
2. Jin, S. 和 Pedersen, T. Duluth urop 在 semeval-2018 任务 2: 通过集成学习和过采样进行多语言表情符号预测。arXiv 预印本 arXiv:1805.10267 (2018)。
3. Wang, Z. 和 Pedersen, T. Umdsub 在 semeval-2018 任务 2: 基于子字嵌入的多语言表情符号预测多通道卷积神经网络。arXiv 预印本 arXiv:1805.10274 (2018)。
4. 吴, C.等人。Thu_ngn at semeval-2018 任务 2: 关注英语表情符号预测的残差 cnn-lstm 网络。第十二届国际语义评估研讨会论文集, 410-414 (2018)。
5. Hoffman, C. 《双语简介第四印象》(Longman Group Ltd, 英国, 1996 年)。
6. Ghosh, S., Priyankar, A., Ekbal, A. 和 Bhattacharyya, P. 代码混合印度英语数据中的情感检测和情感识别的多任务处理。基于知识的系统260, 110182 (2023)。
7. Borth, D., Ji, R., Chen, T., Breuel, T. 和 Chang, S.-F. 使用形容词名词对的大规模视觉情感本体和检测器。第 21 届 ACM 国际多媒体会议记录, 223-232 (2013)。
8. Mohammad, S. M., Sobhani, P. 和 Kiritchenko, S. 推文中的立场和情绪。ACM 翻译。互联网技术。(TOIT) 17, 1-23 (2017)。
9. Majumder, N., Poria, S., Gelbukh, A. 和 Cambria, E. 基于深度学习的文档建模, 用于从文本中进行个性检测。IEEE 英特尔。系统。32, 74-79 (2017)。
10. Chen, T., Xu, R., He, Y. & Wang, X. 使用 bilstm-crf 和 cnn 通过句子类型分类改进情感分析。专家系统。应用。72, 221-230 (2017)。
11. Ghosh, S., Ekbal, A. 和 Bhattacharyya, P. 深度级联多任务框架, 用于检测自杀笔记中的时间方向、情绪和情绪。科学。报告 12, 1-16 (2022)。
12. 帕特瓦, P.等人。Semeval-2020 任务 9: 代码混合推文的情感分析概述。arXiv 电子印刷品 arXiv-2008 (2020)。
13. Cowie, R.等人。人机交互中的情感识别。IEEE 信号处理。马格。18, 32-80 (2001)。
14. Eisner, B., Rocktäschel, T., Augenstein, I., Bošnjak, M. & Riedel, S. emoji2vec: 从描述中学习表情符号表示。arXiv 预印本 arXiv:1609.08359 (2016)。
15. Felbo, B., Mislove, A., Søgaard, A., Rahwan, I. 和 Lehmann, S. 使用数百万个表情符号出现来学习任何领域的表示, 以检测情绪、情绪和讽刺。arXiv preprintarXiv:1708.00524 (2017)。
16. Zhou, X. & Wang, W. Y. Mojitalk: 大规模产生情绪反应。arXiv 预印本 arXiv:1711.04090 (2017)。
17. Barbieri, F., Ballesteros, M. 和 Saggion, H. 表情符号可以预测吗? arXiv 预印本 arXiv:1702.07285 (2017)。

18. Guibon, G.、Ochs, M. 和 Bellot, P. 私人即时消息中的表情符号推荐。第 33 届 ACM 应用计算研讨会论文集, 1821-1823 (2018 年)。

19. Wijeratne, S.、Balasuriya, L.、Sheth, A. 和 Doran, D. Emojinet: 用于表情符号感知发现的开放服务和 API。第十一届国际 AAAI 网络和社交媒体会议 (2017 年)。

20. Santhanam, S.、Srinivasan, V.、Glass, S. 和 Shaikh, S. 我与你们站在一起: 使用表情符号研究危机事件中的团结。arXiv 预印本 arXiv:1907.08326 (2019)。

21. Hussien, W.、Al-Ayyoub, M.、Tashtoush, Y. 和 Al-Kabi, M. 关于使用表情符号训练情感分类器。arXiv 预印本 arXiv: 1902.08906 (2019)。

22. Al-Halah, Z.、Aitken, A.、Shi, W. 和 Caballero, J. Smile, be happy:) 用于视觉情感分析的表情符号嵌入。IEEE 国际计算机视觉研讨会论文集, 0-0 (2019)。

23. Chen, Y.、Yuan, J.、You, Q. & Luo, J. 通过双感表情符号嵌入和基于注意力的 lstm 进行 Twitter 情感分析。第 26 届 ACM 国际多媒体会议论文集, 117-125 (2018)。

24. Ma, W.、Liu, R.、Wang, L. 和 Vosoughi, S. 表情符号预测: 扩展和基准测试。arXiv 预印本 arXiv:2007.07389 (2020)。

25. 卡德尔, F. B. 等人。“当语言失败时, 表情符号占上风”: 一种利用价态反转和语义不一致生成带有表情符号的讽刺句子的新颖架构。计算语言学协会第 61 届年会论文集 (第 4 卷: 学生研究研讨会), 334-351 (2023)。

26. Chakravarthi, B.R.等人。Dravidiancodemix: 代码混合文本中德拉威语言的情感分析和攻击性语言识别数据集。arXiv 预印本 arXiv: 2106.09460 (2021)。

27. Mohbey, K. K.、Meena, G.、Kumar, S. 和 Lokesh, K. 一种基于 cnn-lstm 的混合深度学习方法, 用于猴痘推文的情感分析。新一代计算 1-19 (2023)。

28. Yadav, S. & Chakraborty, T. 代码混合数据的无监督情感分析。arXiv 预印本 arXiv:2001.11384 (2020)。

29. 张, Y.等人。M3gat: 用于会话情感分析和情感识别的多模态多任务交互式图注意网络。ACM 翻译。信息。系统。(2023)。

30. Wang, Z.、Lee, S. Y. M.、Li, S. & Zhou, G. 使用联合因子图模型进行语码转换文本的情感分析。IEEE/ACM 传输。音频语音语言。过程。25, 469-480 (2016)。

31. Wang, Z.、Zhang, Y.、Lee, S.、Li, S. & Zhou, G. 用于代码转换情绪预测的双语注意网络。COLING 2016 论文集, 第 26 届国际计算语言学会议: 技术论文, 1624-1634 (2016)。

32. 肖, L.等人。用于基于多模态方面的情感分析的跨模态细粒度对齐和融合网络。信息。过程。马纳格。60, 103508 (2023)。

33. Mao, R.、Liu, Q.、He, K.、Li, W. & Cambria, E. 预训练语言模型的偏差: 基于提示的情感分析和情绪检测的实证研究。IEEE 传输。影响。计算。(2022)。

34. Singh, G. V.、Firdaus, M.、Ekbali, A. 和 Bhattacharyya, P. 多样性中的统一: 推文中的多标签表情符号识别。IEEE 传输。计算。苏克。系统。(2022)。

35. He, K.、Mao, R.、Gong, T.、Li, C. & Cambria, E. 基于元的自我训练和基于方面的情感分析的重新加权。IEEE 传输。影响。计算。(2022)。

36. Mao, R. & Li, X. 使用基于方面的情感分析和顺序隐喻识别的门控机制桥接多任务学习塔。AAAI 人工智能会议记录 35, 13534-13542 (2021)。

37. Meena, G.、Mohbey, K. K.、Kumar, S. 和 Lokesh, K. 一种用于检测猴痘推文上的情感极性和知识图表示的混合深度学习方法。决定。肛门。J.7, 100243 (2023)。

38. Singh, G. V.、Chauhan, D. S.、Firdaus, M.、Ekbali, A. 和 Bhattacharyya, P. 表情符号、情绪和情感是朋友吗? 针对表情符号、情绪和情绪分析的多任务学习。第 36 届亚太语言、信息和计算会议论文集, 166-174 (2022 年)。

39. Barbieri, F.、Marujo, L.、Karuturi, P.、Brendel, W. 和 Saggion, H. 通过时间变化镜头探索表情符号的使用和预测。arXiv 预印本 arXiv:1805.00731 (2018)。

40. Cappallo, S.、Svetlichnaya, S.、Garrigues, P.、Mensink, T. 和 Snoek, C. G. 新模式: 表情符号在预测、预期和检索方面的挑战。IEEE 传输。多媒体。21, 402-415 (2018)。

41. Lin, W.、Chao, T.、Wu, J. & Su, T. 通过检索策略预测表情符号组合。arXiv 预印本 arXiv:1908.07761 (2019)。

42. Shoeb, A. 和 de Melo, G. 表情符号是否情绪化? 一项旨在了解表情符号与情绪之间关联的研究。arXiv 预印本 arXiv:2005.00693 (2020)。

43. Novak, P. K.、Smailović, J.、Sluban, B. 和 Mozetič, I. 表情符号的情绪。PloS One10 (2015)。

44. Pavalanathan, U. 和 Eisenstein, J. Twitter 上的表情符号与表情符号: 一种因果推理方法。arXiv 预印本 arXiv:1510.08480 (2015)。

45. Chaudhary, A.、Hayati, S. A.、Otani, N. 和 Black, A. W. 多么阳光明媚的一天: 走向表情符号敏感的讽刺检测。W-NUT 2019, 212 (2019)。

46. Corazza, M.、Menini, S.、Cabrio, E.、Tonelli, S. 和 Villata, S. 用于零样本滥用语言检测的基于混合表情符号的屏蔽语言模型。ACL 的调查结果: EMNLP 2020 (2020)。

47. Barbieri, F.、Anke, L. E.、Camacho-Collados, J.、Schockaert, S. 和 Saggion, H. 通过标签关注 lstms 进行可解释的表情符号预测。2018 年自然语言处理经验方法会议论文集, 4766-4771 (2018)。

48. Singh, A.、Blanco, E. 和 Jin, W. 合并表情符号描述可改进推文分类。计算语言学协会北美分会 2019 年会议记录: 人类语言技术, 第 1 卷 (长论文和短论文), 2096-2101 (2019)。

49. 张, Y.等人。学习对话中多模式讽刺检测和情感分析的多任务共性和独特性。IEEE 人工智能汇刊 (2023)。

50. 张, Y.等人。用于对话中多模式讽刺、情感和情感识别的多任务学习模型。信息。富斯。93, 282-301 (2023)。

51. Ilyas, A.、Shahzad, K. 和 Kamran Malik, M. 代码混合罗马乌尔都语-英语文本中的情绪检测。ACM 翻译。亚洲低资源语言。信息。流程。22, 1-28 (2023)。

52. 张, Y.等人。元数据诱导的零样本多标签文本分类对比学习。2022 年 ACM 网络会议论文集, 3162-3173 (2022)。

53. Tesfagergish, S. G.、Kapočiūtė-Dzikienė, J. 和 Damaševičius, R. 使用句子转换器和集成学习进行半监督情感分析的零样本情感检测。应用。科学。12, 8662 (2022)。

54. Hayati, S. A. 和 Muis, A. O. 分析表情符号预测中的情感结合。第十届主观性、情感和社交媒体分析计算方法研讨会论文集, 91-99 (2019)。

55. Kurniawan, S.、Budi, I. 和 Ibrohim, M. O. Ir3218-ui 在 semeval-2020 任务 12: 表情符号对攻击性语言识别的影响。第十四届语义评估研讨会论文集, 1998-2005 (2020)。

56. 乔尔泰金, Ç. & Rama, T. Tübingen-oslo 在 Semeval-2018 任务 2 上: Svm 在表情符号预测方面比 RNN 表现更好。第十二届国际语义评估研讨会论文集, 34-38 (2018)。

57. Liu, M. Emonlp 在 semeval-2018 任务 2: 使用梯度提升回归树方法和双向 lstm 进行英语表情符号预测。第十二届国际语义评估研讨会论文集, 390-394 (2018)。

58. Alexa, L.、Lorent, A. B.、Gifu, D. 和 Traandabat, D. Semeval-2018 任务 2 中的新手: 多语言表情符号预测。第十二届国际语义评估研讨会论文集, 405-409 (2018)。

59. Park, J. H., Xu, P. 和 Fung, P. Plusemo2vec 在 semeval-2018 任务 1: 利用表情符号和 # 主题标签中的情感知识。 arXiv 预印本 arXiv:1804.08280 (2018)。

60. Chen, J., Yang, D., Li, X., Chen, W. 和 Wang, T. Peperomia 在 semeval-2018 任务 2: 基于向量相似度的表情符号预测方法。第十二届国际语义评估研讨会论文集, 428-432 (2018)。

61. Choudhary, N., Singh, R., Rao, V. A. 和 Shrivastava, M. 用于情感分析和多语言表情符号预测的资源稀缺语言 Twitter 语料库。第 27 届国际计算语言学会议论文集, 1570–1577 (2018)。

62. Wang, Z. & Ji, H. 开放词汇脑电图到文本解码和零样本情感分类。第三十六届 AAAI 人工智能会议、AAAI 2022、第三十四届人工智能创新应用会议、IAAI 2022、第十二届人工智能教育进展研讨会、EAAI 2022 虚拟活动, 2022 年 2 月 22 日至 3 月 1 日, 5350–5358 (AAAI 出版社, 2022 年)。

63. Lupart, S., Favre, B., Nikoulina, V. 和 Ait-Mokhtar, S. covid-19 大流行背景下生物医学文章的零样本和少样本分类。 arXiv 预印本 arXiv:2201.03017 (2022)。

64. 伊堀, M.等人。使用切换标记对多个语音文本样式转换任务进行零样本联合建模。 arXiv 预印本 arXiv:2106.12131 (2021)。

65. Olah, J., Baruah, S., Bose, D. 和 Narayanan, S. 使用少量镜头知识迁移进行跨域情感识别。 arXiv 预印本 arXiv:2110.05021 (2021)。

66. Cappallo, S., Mensink, T. 和 Snoek, C. G. Image2emoji: 视觉媒体的零样本表情符号预测。第 23 届 ACM 国际多媒体会议记录, 1311–1314 (2015)。

67. Al-Halah, Z., Aitken, A., Shi, W. 和 Caballero, J. Smile, be happy:) 用于视觉情感分析的表情符号嵌入。 IEEE/CVF 国际计算机视觉会议 (ICCV) 研讨会论文集 (2019)。

68. Kudo, T. & Richardson, J. Sentencepiece: 用于神经文本处理的简单且独立于语言的子词分词器和去分词器。 arXiv 预印本 arXiv:1808.06226 (2018)。

69. Sennrich, R., Haddow, B. 和 Birch, A. 使用子词单元对稀有词进行神经机器翻译。 arXiv 预印本 arXiv: 1508.07909 (2015)。

70. Kudo, T. 子词正则化: 使用多个候选子词改进神经网络翻译模型。 arXiv 预印本 arXiv:1804.10959 (2018)。

71. Pratapa, A., Choudhury, M. 和 Sitaram, S. 用于代码混合语言处理的词嵌入。2018 年自然语言处理经验方法会议论文集, 3067-3072 (2018)。

72. Chiu, J. P. & Nichols, E. 使用双向 lstm-cnns 进行命名实体识别。跨。副教授。计算。语言学家。4, 357–370 (2016)。

73. Ling, W.等人。寻找形式中的函数: 用于开放词汇表词表示的组合字符模型。 arXiv 预印本 arXiv:1508.02096 (2015)。

74. 彼得斯, M.E.等人。深度语境化的单词表示。 arXiv 预印本 arXiv:1802.05365 (2018)。

75. 西特西尔。使用 tf-idf 确定文档查询中的单词相关性。

76. Wu, Z., Liu, Z., Lin, J., Lin, Y. 和 Han, S. Lite 变压器, 具有长短程关注。 arXiv 预印本 arXiv:2004.11886 (2020)。

77. Lee-Thorp, J., Ainslie, J., Eckstein, I. 和 Ontanon, S. Fnet: 将令牌与傅里叶变换混合。 arXiv 预印本 arXiv:2105.03824 (2021)。

78. 张, B. 和 Sennrich, R. 均方根层归一化。 arXiv 预印本 arXiv:1910.07467 (2019)。

79. 吴, F.等人。简化图卷积网络。国际机器学习会议, 6861–6871 (PMLR, 2019)。

80. Deng, J., Guo, J., Xue, N. 和 Zafeiriou, S. Arcface: 深度人脸识别的附加角度边缘损失。 IEEE/CVF 计算机视觉和模式识别会议论文集, 4690–4699 (2019)。

81. Devlin, J., Chang, M.-W., Lee, K. 和 Toutanova, K. BERT: 用于语言理解的深度双向转换器的预训练。 NAACL-HLT 会议记录 (2019 年)。

82. Ekman, P. 基本情感的论证。认知与情感 169–200 (1992)。

83. Ma, W., Liu, R., Wang, L. & Vosoughi, S. 用于表情符号预测的多分辨率注释。2020 年自然语言处理经验方法 (EMNLP) 会议记录, 6684–6694, <https://doi.org/10.18653/v1/2020.emnlp-main.542> (计算语言学协会, 在线, 2020 年)。

84. Ma, W., Liu, R., Wang, L. & Vosoughi, S. 用于表情符号预测的多分辨率注释。2020 年自然语言处理经验方法 (EMNLP) 会议论文集, 6684–6694 (2020)。

85. Yin, W., Alkhalifa, R. 和 Zubiaga, A. 社交媒体上情绪的表情化: 纵向 Twitter 情绪数据集的收集和分析。 arXiv 预印本 arXiv:2108.13898 (2021)。

86. Malte, A., Bhavsar, P. 和 Rathi, S. Team_swift 在 semeval-2020 任务 9 上: 通过代码混合数据的特定领域预训练培养微小数据专家。第十四届语义评估研讨会论文集, 1310–1315 (2020)。

87. Zaharia, G.-E., Vlad, G.-A., Cercel, D.-C., Rebedea, T. 和 Chiru, C.-G. Upb 在 Semeval-2020 任务 9: 使用 Transformer 和多任务学习识别混合社交媒体文本中的情绪。 arXiv preprint arXiv:2009.02780 (2020)。

88. Kingma, D. & Ba, J. Adam: 一种随机优化方法。 arXiv 预印本 arXiv: 1412.6980 (2014)。

89. Rogers, D. J. & Tanimoto, T. T. 用于植物分类的计算机程序。科学 132, 1115–1118 (1960)。

90. Schapire, R. E. 和 Singer, Y. 使用置信度预测改进了增强算法。马赫。学习。37, 297–336 (1999)。

91. Manning, C. D., Raghavan, P. 和 Schütze, H. 《信息检索简介》(剑桥大学出版社, 2008 年)。

92. Himabindu, G. S. S. N., Rao, R. & Sethia, D. 用于代码混合语言的自注意力混合表情符号预测模型: (hinglish)。苏克。网络。肛门。分钟。12, 137 (2022)。

93. 帕特瓦, P.等人。SemEval-2020 任务 9: 代码混合推文的情感分析概述。在 Herbelot, A. 等人中。(编)第十四届语义评估研讨会论文集, 774-790, <https://doi.org/10.18653/v1/2020.semeval-1.100> (国际计算语言学委员会, 巴塞罗那(在线), 2020)。

致谢

作者对印度政府内政部主办的“IITP 针对妇女和儿童的网络犯罪预防卓越中心 - 基于人工智能的妇女和儿童安全工具”的支持表示衷心感谢。我们感谢 MHA 为促进我们的研究和代码混合情绪模型的开发提供的帮助。

作者贡献

G.V.S.: 方法、实施、资源创建、实验; S.G.: 方法论、实施、资源创建、撰写手稿; M.F.: 资源创建、方法论、撰写手稿; A.E: 概念化、方法论、草案审查、资金获取、监督; P.B.: 概念化、监督、资金获取。所有作者均已阅读并同意稿件的出版版本。

利益竞争

作者声明没有竞争利益。

附加信息

信件和材料请求应发送至 G.V.S.

重印和许可信息可在 www.nature.com/reprints 上获取。

出版商说明施普林格·自然对于已出版地图和机构隶属关系中的管辖权主张保持中立。

开放获取本文根据知识共享署名 4.0 国际许可证获得许可，该许可证允许以任何媒介或格式使用、共享、改编、分发和复制，只要您对原作者和来源给予适当的认可，提供知识共享许可证的链接，并指出是否进行了更改。本文的图像或其他第三方材料包含在文章的知识共享许可中，除非材料的信用额度中另有说明。如果文章的知识共享许可中未包含材料，并且您的预期用途不受法律法规允许或超出了允许的用途，您将需要直接获得版权所有者的许可。要查看此许可证的副本，请访问 <http://creativecommons.org/licenses/by/4.0/>。

© 作者 2024