叙事文本的无监督情绪识别框架的系统评估

萨米拉·扎德（Samira Zad）和马克·芬利森（Mark A.Finlayson）

计算与信息科学学院

佛罗里达国际大学11200 SW 8th St.，Miami，FL 33199，USA

{ szad001，markaf } @ fiu.edu

摘要

识别文字表达的情感

（又名文字情感识别）在过去十年中受到了广泛关注。叙事通常涉及大量的情感表达，因此叙事文本上的情感识别对于叙事理解的计算方法引起了极大的兴趣。Kim等人的先前工作。（2010年）是童话语料库中报道的情感检测性能最高的作品。但是，仔细检查该工作发现了严重的可重复性问题，我们无法重新实现Kim所述的方法。因此，我们实施了一个由Kim的方法启发的框架，在该框架中，我们仔细评估了主要的设计选择。我们确定了表现最高的组合，其表现比Kim的报告业绩高7.6*F* 1分。对带注释的数据进行仔细检查后发现，相关词典WordNetAffect（WNA; Strapparava and Valitutti ，2004 ）中存在大量缺失和不正确的情感术语，这使我们能够以有用的方式对其进行扩充。更一般地说，这表明WNA缺少许多明显具有情感意义的单词和短语，这表明在增加或改进情感本体方面投入的精力可能对改善情感识别系统的性能有用。我们发布我们的代码和数据，以确保将来能够重做这项工作。

# 引言

情感是交流的主要方面，可以通过多种方式进行传递，包括手势，面部表情，语音和文本。由于这种重要性，自动情感识别对于许多应用程序都非常有用，包括自动叙事理解。叙事是“具有关联性的事件和角色的表示，具有可识别的结构，在空间和时间上受限制，并且包含有关所解决主题的隐式或显式消息”（Kreuter等人，2007 ，第222页），叙事通常用于表达作者和角色的情感，并在观众中引起情感。对于许多叙事而言，只需要考虑*罗密欧与朱丽叶等爱情小说*或电影《*泰坦尼克号》*毫不夸张地说，缺乏对情感的理解会导致对叙述意义的严重贫困。

由于感觉情绪和语言表达之间的复杂关系，因此情绪识别是一个具有挑战性的问题。这不仅包括标准的自然语言处理挑战，例如多义词和共指解析的难度（Uzuner等人，2012 ；Peng等人，2019 ），还包括特定于情感的挑战，例如上下文如何巧妙地改变情感解释（Cowie等，2005 ）。由于解决该任务的质量标记数据不足，这些技术挑战更加严峻。

|  |
| --- |
| 26  *第一届叙事理解，故事情节和事件联合研讨会论文集*，2020 年7月9日，第26-37页。2020年计算语言学协会 |

在情感识别方面已有许多先验工作。关于叙事，金等人。（2010年）报道了一种高性能的童话语料库上的情感识别方法（Alm ，2008年）。这种方法涉及一种无监督的学习框架，用于文本数据中的情感识别，它使用了埃克曼情绪心理学理论的一种改进形式（欢乐，愤怒，恐惧，悲伤；埃克曼，1992b ）。在这项工作中，他们使用了WordNetAffect（WNA）和ANEW（英语单词的情感规范）情感词典来构建语义空间。每个句子都使用*tf-idf*放置在空格中在词典中找到情感词的权重。然后，他们测试了三种方法-非负矩阵分解（NMF），潜在语义分析（LSA），概率概率潜在语义分析（pLSA）-压缩空间以提取构造的向量空间模型的特征，减少噪声并消除异常值。最终，该框架使用余弦相似度来标记句子，方法是评估它们与基于与情感词典（更具体而言是WNA的扩展）紧密相关的WNA条目生成的标准向量的相似度，从而将其相比较。表现最好的方法是NMF，他们报告称该方法的平均情绪识别*F* 1为0.733。

然而，对工作的仔细检查显示出严重的可重复性问题。尽管尽了最大的努力，我们仍无法在Kim所报告的表现附近的任何地方再现结果。实际上，我们的最佳尝试仅产生了大约0.25 *F* 1。这是由于几个原因。首先，本文缺乏有关模型超参数的信息。其次，本文省略了NMF关键步骤的描述，包括如何识别代表性特征以及在语义空间压缩之前应删除哪些特征。第三，本文没有解释如何使NMF适应文本NMF模型中出现的稀疏矩阵。第四，与WNA相关的某些资源或者没有正确识别，或者不再可用。这些遗漏使我们无法以任何精确度复制他们的模型。

因此，我们承诺对Kim等人所述的设计空间进行系统的探索。（2010 ）。我们研究了由以下人员报告的性能最高的向量空间压缩技术

Kim*等。*（NMF）以及主要成分

分析（PCA）和潜在Dirchelet分配

（LDA）在其他工作中被报告为高性能技术。我们表明NMF确实表现最好，并且我们清楚地说明了我们的实验设置，包括用于识别相关特征和处理稀疏文本矩阵的方法。本文实现的PCA和NMF方法是基于Mairal等人的工作。（2009 ）以及Boutsidis和Gallopoulos （2008 ）分别实施了适用于大型稀疏矩阵的机制（在我们的案例中，

1 *，* 090×2 *，* 405 ）。这项工作使*F* 1的性能比Kim报告的结果提高了约7.6点。我们发布代码和数据以方便将来的工作[[1]](#footnote-1)[1]。[[2]](#footnote-2)

本文其余部分的结构如下。我们简要回顾了情绪的心理模型，描述了几种重要的情绪语言资源，并概述了许多著名的情绪识别模型（第2节）。然后，我们描述我们的自适应无监督情绪识别方法，并详细描述所有步骤，参数和所需的资源（第3节）。接下来，我们将描述我们的方法在Alm的童话语料库（Alm ，2008 ）上的执行情况，该方法在句子层面上对情感进行了注释（第4节）。最后，我们确定了一些未解决的挑战，这些挑战指向未来的工作（第5节），并总结了我们的贡献

（第6节）。

# 相关工作

2.1心理情感理论情感理论可以追溯到古希腊人和罗马人，并且一直是人们对历史上人类经验的本质进行探究的主题，其中包括19世纪查尔斯·达尔文和威廉·詹姆斯提出的著名建议（达尔文和Prodger ，1998年；James ，1890年）。现代的情绪心理学理论可以分为两类：*范畴的*和*维度的*（Calvo and Mac Kim ，2013 ）。分类心理学模型提出了离散的基本情绪，例如，Oatley和Johnson-Laird（1987年））包含五个基本情感，几个具有六个基本情感的模型（Ekman ，1992b ; Shaver et al。，1987 ），Parrott的六个基本情感模型排列在三层树中（2001 ），Panksepp的七个情感模型（1998 ） ，以及Izard的《十岁》（2007年）。

相比之下，维度心理模型通过将情感定位在可能包含唤醒，化合价，强度等维度的空间（通常为2到4个）中来确定情感。这些维度包括二维模型，例如Russell的绕线模型（1980年），Scherer增强的Circumplex （2005 ）和Whissell的模型（Cambria ，2016 ）。Lovheim模型（¨ 2012年）是一个使用三个维度的示例，而Ortony等人（2012年）则使用了三个维度。（1990 ），Fontaine等人。（2007年），以及Cambria等人。（2012 ）提出了三维模型。

最后，也有它结合了分类和尺寸方面的模型，称为*混合*车型，其中最突出的是普拉奇克的轮体模型用八种基本情绪（普拉奇克，1980年，1984年，2001年）。

在已提出的所有情感模型中，埃克曼的六类模型（愤怒，厌恶，恐惧，幸福，悲伤，惊奇）是迄今为止在计算方法中最受欢迎的模型，部分原因在于其简单性，部分原因在于它成功地应用于面部表情自动识别（Zhang et al。，2018 ; Suttles and Ide ，2013 ; Ekman ，1992b ，a ，1993 ）。尽管有一些研究人员怀疑埃克曼的模型是否完整，因为它似乎嵌入了西方文化偏见（Langroudi等人，2018）。如下所述，在我们自己对情绪识别系统的评论中，Kim等人描述了为叙述文本报告的性能最高的系统。（2010 ）。在这项工作中，他们使用了埃克曼模型的四个标签子集（幸福，愤怒，恐惧和悲伤），这就是我们在本文中采用的模型。

2.2 情感词汇

文本中用于情感识别的主要语言资源之一是情感词典，它只是与情感类别相关的单词列表。情感词典既可以用于基于规则的识别方法，也可以用于基于机器学习的识别方法。有两种类型的情感词典。一种是通用情感词典（GPEL），用于指定情感词的一般含义。GPEL有时会将情感表达为分数，并且可以应用于任何领域。突出的GPEL包括WordNet Affect（WNA；Strapparava和Valitutti ，2004 ），威斯康辛州感知属性数据库（WPARD； Medler等，2005 ），语言查询和字数统计

（LIWC；Pennebaker等，2001 ），以及国家研究委员会（NRC）和NRC Hashtag词典（Mohammad和Turney ，2010 ；Mohammad等，2013 ）。第二种词典是领域特定的情感词典（DSEL），其针对特定领域进行情感识别。Bandhakavi等。（2014 ）提出了针对情感推文的特定领域词典。表1比较了几个关键GPEL的详细信息。

WordNet Affect版本1.1 Kim等。用过的

WordNet Affect（WNA；Strapparava和Valitutti ，2004 ），它建立在通用WordNet数据库的基础上（Fellbaum ，1998 ）。WNA分类280

WordNet*名词* 以情感层次结构为基调，该层次结构植根于Ekman基本情感的增强版本，并在图1中进行了部分描绘。单词-

Net将其他1,191个*动词*，*副词*和*形容词*同义词集链接到此以*名词*为核心的层次结构。这些同义词集代表大约3500个英语lemma-POS对。

2.3情绪识别方法至少有一百篇论文描述了文本中的情绪识别方法

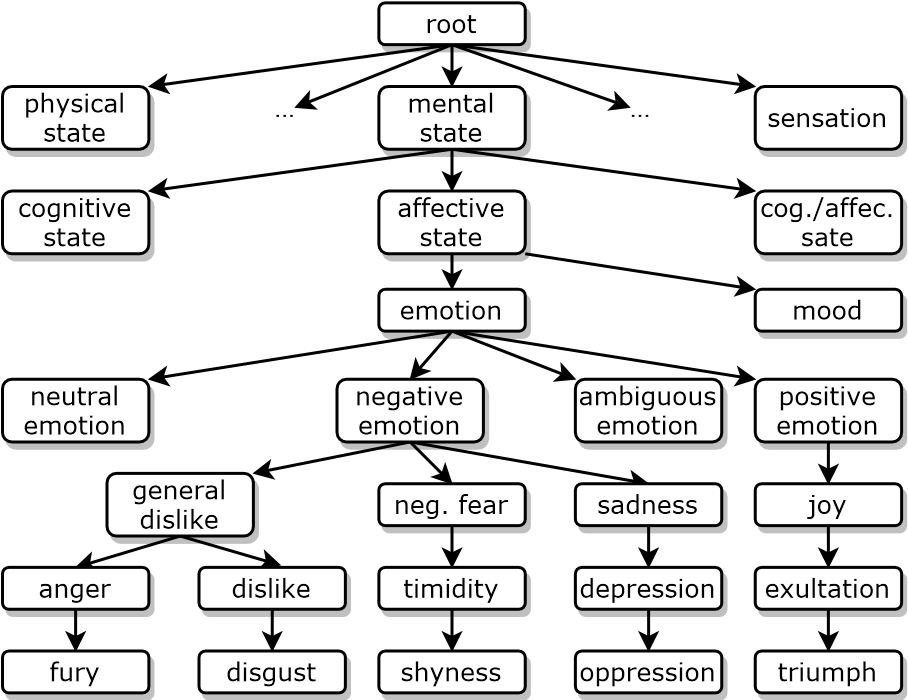


图1：WordNet Affect版本1.1中的情感层次结构。

（Calefato等人，2017 ; Teng等人，2007 ; Shaheen等人，2014 ）。在这里，我们回顾了一些已应用于叙事类或与叙事相关的话语类型的方法的选择。重要的是要记住，所有这些方法都使用不同的数据和不同的理论，通常涉及不同数量的标签。在所有条件都相同的情况下，分类结果通常会随着数字标签的增加而降低。因此，只能粗略地比较每个系统的性能。

Strapparava和Mihalcea （2008 ）在新闻头条中描述了一种识别情绪的系统。他们从各种新闻网站（例如Google新闻，CNN和在线报纸）中提取了1,250个新闻标题，并使用Ekman的模型（愤怒，厌恶，恐惧，喜悦，悲伤和惊奇）进行注释，将数据分为一组训练250个测试集和1,000个测试集（这称为*SemEval-2007*数据集）。他们测试了五种途径：WNA-P RESENCE ，LSA-S英格尔-

w ^ ORD ， LSA-E MOTION -S YNSET ， LSA-A LL -

è MOTION -W ORDS ，和N AIVE乙者占多-T RAINEDON -B原木。WNA-P RESENCE查找WNA中列出的标题词，提供了0.38的最佳精度。的LSA-A LL -E MOTION Wˉˉ ORDS ，其计算之间的六个影响字和标题的LSA表示，导致最高的召回和矢量相似度*˚F* 1，分别在0.90和0.176。

Aman和Szpakowicz （2008 ）使用了支持

向量机（SVM）在博客数据上进行了培训和测试，以识别Ekman的情绪类别，另外还增加了两个类别： *混合情绪*和*无情绪*。由自动检索的博客中的四位人类法官手动注释了1,890个句子，以创造出

|  |
| --- |
| 表1：与情感相关的词典表。WNA = WordNet影响；NRC =加拿大国家研究委员会；LIWC =语言查询和字数统计；WPARD =威斯康星州感知属性评级数据库；ANEW = |

|  |  |  |  |
| --- | --- | --- | --- |
| 情感词典 | 引文 | 一套情绪 | 参赛作品 |
| WNA | Strapparava和Valitutti （2004 ） | 情绪层次 | 915个同义词集 |
| NRC / Emolex | 穆罕默德与特尼（2010 ） | Plutchik基本模型1980 ，neg。/ pos。 | 14,182 |
| 生命力中心 | Pennebaker等。（2001年） | 是否有效，否定/正负。焦虑，愤怒，悲伤 | 5,690 |
| NRC标签 | Mohammad等。（2013年） | Plutchik的基本模型 | 32,400 |
| WPARD | 梅德勒（Medler）等人。（2005年） | 正面或负面 | 1,402 |
| 重新 | 布拉德利和郎（1999 ） | 3D（价，唤醒，主导） | 1,035 |

英语单词的情感规范

吃了语料库。SVM的功能是Roget词库和WNA中列出的情感词。每种情感等级的*F* 1测度范围在0.493至0.751之间，在每种情况下都超过了基线表现。

德久等。（2008年）描述了一种基于词典的日语情感识别系统。他们通过从日语表达评估（JEE）词典中识别出349个情感词来手工制作情感词典，这些单词分为10种不同的情感：3种正面（幸福，愉悦，放松）和7种负面（恐惧，悲伤，失望，不愉快，孤独，焦虑和愤怒）。然后，他们使用此词典自动组合了带标签的130万个激发情绪（EP）的“事件”的语料库（定义为从属从句，修改了情绪陈述）。然后，他们展示了一种两步式的情感识别方法，首先是基于SVM的粗略情感极性分类（正，负或中性），然后是基于kNN的非中性实例分类为适当的细粒度情感类别（对于积极的情感类别为3） ，7（否定）。他们的最佳表现模型报告的准确度在0.5到0.8之间。

Cherry等。（2012年）提出了两个监督的机器学习模型，用于自杀笔记句子中的情绪识别。他们使用了2011 i2b2 NLP挑战任务2，该任务包含训练集中的4,241个句子和测试集中的1,883个句子，并用13个情感标签进行手动注释。每个情感一个分类器的方法产生的*F* 1为0.55，而每个句子应用多个情感标签的潜在序列模型的*F* 1为0.53。他们指出，超过73％的培训数据缺少标签，这限制了培训的有效性。Bandhakavi等。（2017年）使用unigram混合模型（UMM）进行实验，以识别推文，事件报告，新闻头条和博客中的情绪。每个语料库都用不同的情感理论进行了手动注释：280,000条带有Parrott的六个主要情感的推特（Parrott ，2001年），1,250个新闻标题和5500个带有Ekman的六个情感集合的博客，来自ISEAR数据集的7000个事件报告[[3]](#footnote-3)[2]标记有七个情感。放。该研究的目标之一是比较特定领域情感词典和通用情感词典（DSEL与GPEL）的效用。他们发现，将DSEL词典词与n-gram，语音标签的一部分以及情感词典中的其他词结合在一起，可以得到0.60 *F* 1的最高性能。[[4]](#footnote-4) 在博客数据上。

Kim等。（2010年）报告了叙事文本中性能最高的情感识别系统。在他们的数据中，有一组176个童话故事，

150万个句子被Alm （2008 ）标记为埃克曼理论的四个情感子集（愤怒，恐惧，喜悦和悲伤）。他们展示了一种无监督的方法，其中每个句子在情感词空间（从WNA和ANEW提取）中被转换为向量，然后使用降维技术（NMF，LSA或pLSA）进行压缩。然后将这些向量与在相同空间中针对四种情绪中的每一种计算出的参考向量进行比较。他们报告表现为*F* 1NMF为0.733，这是他们表现最好的模型。这种方法的一个优势是它不受监督，这意味着既不需要大量的训练数据，又可以将所有带注释的数据用于测试。这很重要，因为要测试该技术的语料库很小。

# 情绪识别框架

现在，我们描述一个基于Kim报告的情感模型的无监督系统

|  |
| --- |
| 表2：类似叙事文字的情感识别方法，按表现排序。LSA =潜在语义 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 引文 | 语料库 | 词典 | ＃情绪 | 方法 | *˚F* 1 |
| Kim等。（2010年） | 童话故事 | WNA | 4 | NMF | 0.73 |
| Bandhakavi等。（2017年） | 鸣叫 | UMM + DSEL | 6 | 仅词典 | 0.64 |
| 阿曼（Aman）和斯帕科维奇（Szpakowicz）（2008 ） | 博客 | -- | 6 | 字母组合 | 0.57 |
| Cherry等。（2012年） | 自杀笔记 | -- | 15 | 支持向量机 | 0.55 |
| Strapparava和Mihalcea （2008 ） | 头条新闻 | -- | 6 | LSA | 0.17 |
| 德久等。（2008年） | “ EP”活动 | JEE Dict。 | 10 | 支持向量机 | 0.5–0.8 Acc。 |

分析; LS =潜在序列建模

等。（2010 ）。在遵循该工作的一般模式的同时，我们尝试了另一套降维方法（Lee和Seung的NMF以及PCA和LDA）。系统输入以下各项作为输入：

* 包含*n个*句子*S的*语料库：*s* 1 *，s* 2 *，*

*...，s n* ;

* 一组情绪*E* = { *e* 1 *，e* 2 *，...，e l* −1 *，*中性}，用于将情绪分为*l个*不同的类别，包括中性；和，
* 情感词库*大号*：Ω7→ *Ë*其中每个字映射在语料库*ω* ∈Ω （其中Ω 具有*米*计）与一种情绪*Ë* ∈ *Ë*。单词*ω*具有其词素化形式，并具有特定的POS。

该系统的流程图如图2所示。该系统包括四个连续的步骤。在第一步（预处理）中，系统使用CoreNLP库处理输入语料库（Manning等，2014 ），将文本分为句子和定形标记。第二步，向量空间建模，使用经过去势化的标记为向量空间中每个句子生成一个向量，该向量的维数对应于Ω中的项。第三步，噪声消除或降维，我们探索了三种不同的模型（非负矩阵分解，潜在狄利克雷分配和主成分分析）以减小维数或提取向量空间的特征。我们在此所做的主要贡献之一是分析并解释了此步骤对最终情绪识别系统性能的影响。最后，标记的第四步将每个句子的向量与每种情感的向量进行比较，选择最接近的情感作为句子的标签。

增强WNA 如前所述，WNA

1.1将情感标签分配给WordNet的1,471个同义词集（同义词集）。这对应于将近3,495个情感引理-POS对的词典。对WNA的仔细检查发现，两者

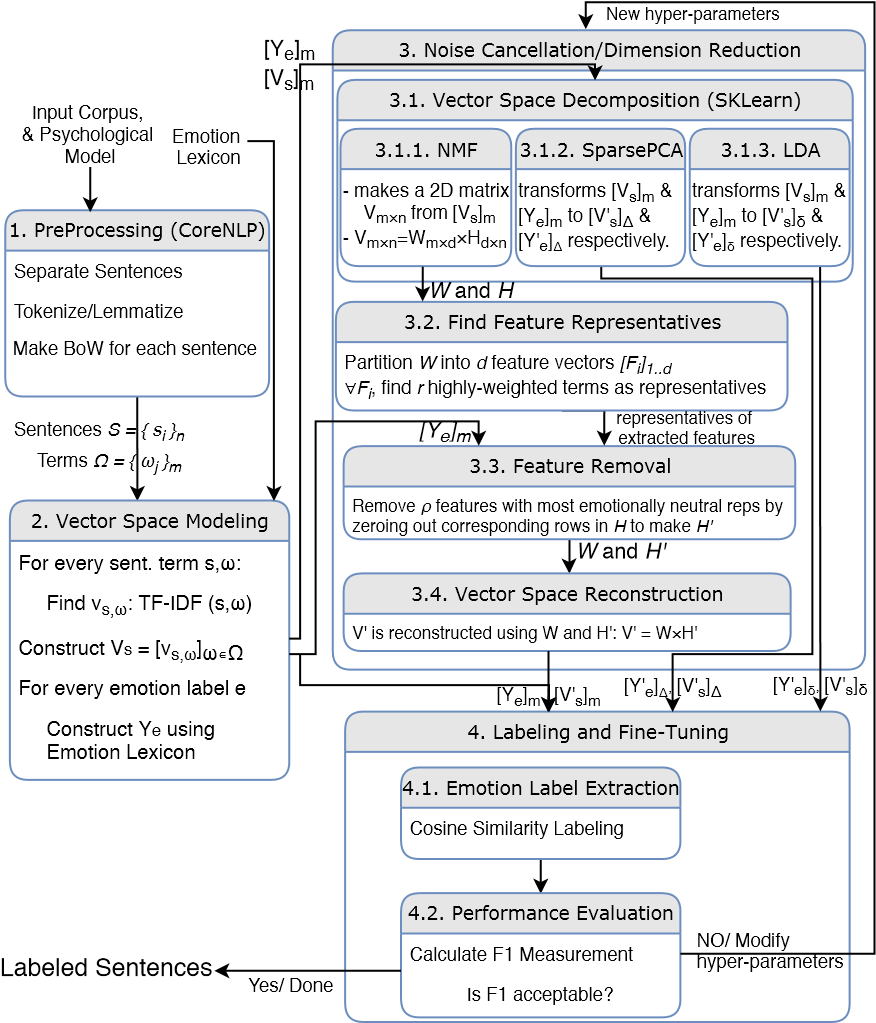


图2：拟议系统的流程图。[ *V s* ] *m*和[ *Y e* ] *m分别*表示原始的*m*维句子和情感矢量模型， [ *V s* 0 ] ∆ 和[ *V s* 0 ] *δ*表示使用NMF，PCA和LDA的转换后的句子矢量模型技术分别。和分别使用PCA和LDA技术表示经变换的情感矢量模型。 

正确地包括以及缺少对。对于错误地包含的对，包括了相当数量的对，因为它们的所有多种感官都以与次要情感感官（而不是其主要的非情感感官）有关的情感来标记。我们手动检查并删除了这些不正确的标签。此外，借助WNA已标记的紧密相关对，我们识别出缺失的lemmaPOS对。例如，WNA中缺少*glorious-JJ*对，但是（通过*派生*关系）与已经标记的对*glorify-VB相关*。我们手动搜索这些缺失的关系，添加缺失的术语，以及递归地添加它们的同义词（例如，*光荣的-JJ*导致*出色*， *壮丽*，*辉煌*和*精湛*）。我们总共删除了613个标签，并添加了814个不同lemma-POS对的标签，最终计数为4048个lemma-POS对。

通常，使用情感术语的固定词典来捕获高度依赖于上下文的情感表达的技术充其量是有问题的。尽管我们在此处显示了改进词典的工作确实可以改善情感识别结果，但是最终，任何技术都必须从基于刚性词典的方法转变为更灵活的方法。我们计划在未来的工作中探索这种方向。

步骤1：预处理

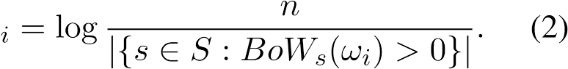
对于每一个句子*小号*∈*小号*在给定的语料库，我们通过令牌化的句子，lemmatizing每个单词构建词袋。我们生成用于弓计数矢量*小号* 通过在句子中的每一个映射到引理计数（Ω7→ Ž ≥ 0）。我们不会删除停用词，因为下一步将通过*tf-idf*计算将其影响降至最低。

步骤2：向量空间建模

使用在第一步骤构成的计数矢量，我们计算一个*TF-IDF*每个句子以及为每种情感类别标准矢量矢量*Ë* ∈ *Ë*。对于每个句子*小号Ĵ* ∈*小号*，我们构建*米*维向量，其中向量中的每个条目是*TF-IDF*术语的*ω我*在句子*小号Ĵ* ; IE

*v ij* = TF *i，j* × IDF *i* （1）

其中TF *I，J* =的BoW*小号Ĵ*（*ω我*），

以色列国防军

*n*是句子的数量，而。

构造的向量空间模型由以下*m* × *n*矩阵 *V表示* ：

 *v* 1 *Ĵ* 

 *v* 2 *Ĵ* 

*V* = [ *V小号*1个*V小号*2 *... V小号Ñ* ]其中*V小号Ĵ* = ..。 （3）



*v MJ*我们计算每个情感类别标准向量*Y Ë* =（*Y E，ω* 1 *，Y E，ω* 2 *，...，Y E，ω米*），其中*Y E，ω我*是1，如果术语*ω我*被词典映射到*e*，否则为0。

步骤3：消除杂讯或减少尺寸

上一步中的向量*V s*和*Y e*均为*m*维向量，其中*m*是语料库中项的总数。有许多术语对句子的情感标签几乎没有影响。因此，降维或噪声消除技术可以改善随后出现的情绪标记步骤的性能。主成分分析（PCA）可以消除噪声已有相当长的一段时间（Abdi and Williams ，2010 ），而潜在狄利克雷分配（LDA）是专门为减少自然语言处理中的维度而开发的（Blei等人，2003）。）。Lee和Seung （1999 ）首次引入非负矩阵分解（NMF）来消除噪声。

步骤3.1：向量空间分解

我们可以通过以下三种方式之一来分解获得的矩阵*V*：

1. 非负矩阵分解（NMF）：我们使用NMF从句子的*m*维向量中提取*d个*特征。
2. 主成分分析（PCA）：我们将*V s*向量的维数从*m减少*到∆ *<m*。
3. 潜在狄利克雷分配（LDA）：我们将*V s*向量的维数从*m减少*到*δ<m*。

当使用PCA或LDA时，我们可以直接移至系统的第四步。但是，对于NMF，我们必须选择重要的项（步骤3.2），删除不相关的特征（步骤3.3）并重建向量空间（步骤3.4）。

当使用NMF分解向量空间模型时，将*V*分解为两个矩阵*W m* × *d* = [ *w ij* ]和*H d* × *n* = [ *h ij* ] ，它们都具有所有非负项：

* 1. = *w ^* × *ħ* ST*瓦特IJ* ≥0和*ħ IJ* ≥0 （4）

请注意，在此步骤中，*d*被视为超参数，可以通过在开发集中最大化系统的输出来微调*d*的数值。

NMF因式分解过程产生矩阵*W*，矩阵*d*列中的每一个代表语料库中原始*n个*句子中每个句子的*m*维特征：

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *š* 1 *小号*2个 *小号Ñ ˚F* 1 *˚F* 2 *˚F d*    *V*  = [  *v*  *ij*  ]  *m*  ×  *n*  ↓  *W*  = [  *w*  *ij*  ]  *m*  ×  *d*  ↓  *H*  = [  *h*  *ij*  ]  *d*  ×  *n*  y  *t*  1  →            *v*  11  *v*  12  *v*  *...*  1  *n*  *v*  21  *v*  22  *...*  *v*  2  *n*  .  .  .  .  .  .  .  .  .  .  .  .  *v*  *m*  1  *v*  *m*  2  *...*  *v*  *mn*                      *w*  11  *w*  12  *w*  *...*  1  *d*  *w*  21  *w*  22  *w*  *...*  2  *d*  .  .  .  .  .  .  .  .  .  .  .  .  *w*  *m*  1  *w*  *m*  2  *...*  *w*  *md*                      *h*  11  *h*  12  *h*  *...*  1  *n*  *h*  21  *h*  22  *...*  *h*  2  *n*  .  .  .  .  .  .  .  .  .  .  .  .  *h*  *d*  1  *h*  *d*  2  *h*  *...*  *dn*            ←  *t*  2  →  =  ×  ←  .  .  .  *t*  *m*  →  ←  weight of  *F*  *d*    ↓ ↓ ↓ ↓ ↓ ↓   重量的*˚F* 1重量*˚F* 2  图3：非负矩阵分解（步骤3.1）以提取句子向量模型*V的特征*。该过程的结果由矩阵*W*和*H给出*。*W的*列对应于模型的提取特征*F* 1 *，F* 2 *，...，F d*，*H的*行称为这些特征的权重。   |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | 代表1 →代表2 →  代表3  。 | *F* 1次。  ↓ |  | *˚F我*代表。  ↓ |  | *˚F e*代表。  ↓ |  | ...  *ħ博士*2 *d* | ... *...* |  | | NN的好意  JJ挚爱  NN机会  VB很高兴  NN饮料  ... | *...* | VB飞跃  NN鸡  狮子  VB服从  RB大声  ... | *...* | NN悲伤VB工资  NN悲惨  NN绝望  VB领带  ... |   图4：通过将矩阵*H中*的相应权重清零来删除最不相关的特征。更新后的*H*矩阵由*H* 0表示。句子矢量模型然后被乘以重构*Wˉˉ*由*ħ* 0（步骤3.3＆3.4）。更新后的句子矢量模型由矩阵*V* 0表示。 |

*瓦特*1 *Ĵ* 

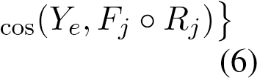
*瓦特*2 *Ĵ* 

* 1. = [ *˚F* 1 *˚F* 2 *˚F* 3 *... F d* ] 其中*˚F Ĵ* =... （5）

*w ^ MJ*

*H*矩阵的*d*行中的每行代表*F中d个*特征的权重。这种分解如图3所示。

步骤3.2：项选择对于每个特征*F j*，我们确定权重最大的项的分数*r*作为其代表，其中*r*是可以在系统优化期间进行微调的超参数（*r*通常小于1％ ）。

步骤3.3：特征去除在这一阶段，我们去除具有很少或没有情感相关性的*ρ*特征，其中*ρ*是可以调整的非负整数超参数。如果该功能的所有代表性术语（如在上一步中选择的那样）被词典标记为中立，则我们将其称为“情感无关”。这些功能将始终被首先删除。如果*ρ*小于情感无关特征的数量，则我们随机选择。另一方面，如果与情感无关的特征的数量少于*ρ*，则消除特征*F j*在他们的整体情绪相关性，这是通过估计余弦相似性比率的情感向量之间的标准偏差来计算的顺序*Y Ë* “在步骤2和S获得*˚F Ĵ* ◦ *ř Ĵ*（逐元素的乘积*˚F Ĵ*和*- [R Ĵ*），其中*R j*是项是否代表*F j*的二进制标识符，并且是根据步骤3.2的结果构造的。象征性的，量化的特征是如何在感情上相关*˚F Ĵ*是，我们计算了以下标准偏差：*σ Ĵ* = STDDEV*Ë* ∈ *Ë* \中性SIM卡

步骤3.4：向量空间重构

在此步骤中，将重建向量空间模型

（*V* 0）消除了不相关的特征之后。让*我*表示一组索引，这些索引的相应特征在上一步中被确定为最不相关。然后，重构的向量空间为：

*V* 0 = [ *v ij* 0 ] *m* × *n* st *v ij* 0 = X *w ik h kj* （7）

1≤ *ķ* ≤ *DK /* ∈*我*

图4说明了向量空间重构。

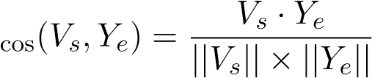
步骤4：贴标签

最后，通过在NMF，PCA或LDA的帮助下测量句子向量*V s*与标准情绪向量*Y e*之间的相似性，进行情感识别过程。每个句子*s的*标签通过以下公式计算：*s的*预测标签= argmax sim （*V s，Y e*） （8）

*Ë* ∈ *Ë*

|  |
| --- |
| 图5：探索NMF的超参数空间。超参数*d*，*r*和*ρ的*每种组合（维度，代表和移除的特征）都会为每个情感标签产生特定的*F* 1分数。与模型（*d，R*）=（975 *，* 10） ，与绿色色彩突出显示，导致最高的整体*˚F* 1得分当*ρ* = 18 。对于每个单独的情感，最好*˚F* 1场得分在（a）中喜悦实测值：（*d，R，ρ*）=（1050 *，* 10 *，* 3），（b）中愤怒：（*d，R，ρ*）=（1025 *，* 3 *，* 24） ，（C）恐惧：（*d，R，ρ*）=（1000 *，* 6 *，* 15） ，（d）悲伤：（*d中，r，ρ*）=（975 *，* 6 *，* 15） 。 |

其中相似度函数可以通过两个给定向量所成的角度的余弦来度量：

SIM卡（9）

# 童话数据的表现

我们使用由Alm （2008 ）构造的人工注释的童话数据集对系统进行了调整和测试，该数据集包含176个儿童的童话故事（Brothers Grimm的80个故事，Hans Andersen的77个故事，Beatrix Potter的19个故事），15,087个独特的句子（15,302个）句子），7,522个独特词和320,521个总词。这些童话故事由两名注释者注释，将每个句子的情感和情绪标记为欢乐，愤怒，恐惧，悲伤或中立，其中每个句子四个标记。在整个句子中，只有1,090个同意*所有四个非中性标签*。Kim等。（2010 ）仅使用这些句子来训练和测试他们的系统[[5]](#footnote-5)[3][[6]](#footnote-6)，我们遵循相同的步骤。有2,405个唯一的term-POS对。另外，在数据集中的标签分布在图6所示的饼图中指定。

40.7

%

Joy

19.9

%

Anger/Disgust

15.1

%

Fear

24.2

%

Sadness

图6：带有注释者之间一致同意的句子的童话标签分布。

我们根据Alm的数据评估了系统的性能。在不增加WNA的情况下，使用WNA的原始1,471个同义词集，*F* 1得分为0.625。表4所示的性能指标是通过使用增强WNA的模型获得的。图5所示的图显示了使用NMF技术消除噪声的提议模型的各种设置的*F* 1分数。另外，表4总结了我们的系统针对四个情感类别中每个类别的精确度，召回率和*F* 1得分，以及其总体*F* 1使用NMF，PCA或LDA和不同设置（超参数值）时的得分。如该表观察到的，最高的总体*˚F* 1使用时，获得得分

NMF与（*d，R，ρ*）=（975 *，* 10 *，* 18） 。在该模型中，有209个句子被错误地标记。其中，一些具有挑战性的例子在表3中。

# 未解决的挑战和未来的工作

如已经讨论的，关于自动情感识别的一个挑战是情感语义的上下文相关性。例如，*我在月球上！*是一种极端幸福的表达，但不使用任何明确的幸福或快乐的词（或者根本不使用任何情感词）。另一个障碍是多义词，即单词既具有情感意义又具有非情感意义。识别正在使用的单词含义是具有挑战性的，并且仍然是一个悬而未决的问题。除了这些基本问题外，还严重缺乏高质量的注释数据，不仅用于叙述文本，而且对于所有语篇类型都如此。带注解的语料库使用了多种有时互不兼容的情感理论，并且通常注解不佳，且注解者之间的协议较少且存在许多错误。

考虑到这些考虑因素，未来的工作有许多可能的方向，例如：

* 调和情感词典和上下文去

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | 句子 | 预料到的 | 金标 | | *他们告诉他，父亲病得很重，他们担心没有什么能救他。* | 害怕 | 悲伤感 | | *并在桥梁的视线中！说可怜的小猪，几乎哭了。* | 悲伤感 | 害怕 | | *她再次微笑，然后人们说她死了。* | 悲伤感 | 喜悦 | | *然后，他猛地瞄准，击中了狼的头，并当场杀死了他！* | 愤怒 | 喜悦 |   *……当他死后，他们割开了身体，让汤米获得了自由。*  表3：具有最准确设置的模型错误标记的句子的具有挑战性的示例。   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | NMF | 1025、3、24  1000,6,15 | 0.859  0.872 | 0.876  0.858 | 0.867  0.865 | 0.884  0.861 | 0.705  0.687 | 0.785 0.764 | 0.682  0.692 | 0.715  0.764 | 0.698  0.726 | 0.733  0.742 | 0.799  0.830 | 0.764  0.784 | 0.800  0.804 | 0.799  0.803 | |  | 975,6,15 | 0.860 | 0.874 | 0.867 | 0.882 | 0.691 | 0.775 | 0.689 | 0.739 | 0.713 | 0.759 | 0.833 | 0.794 | 0.808 | 0.807 | |  | 975,10,18 | 0.858 | 0.874 | 0.866 | 0.879 | 0.705 | 0.783 | 0.703 | 0.733 | 0.718 | 0.755 | 0.830 | 0.791 | 0.809 | 0.808 | |  | 1150 | 0.885 | 0.764 | 0.820 | 0.743 | 0.719 | 0.731 | 0.542 | 0.745 | 0.628 | 0.748 | 0.765 | 0.757 | 0.752 | 0.683 | |  | 950 | 0.883 | 0.766 | 0.820 | 0.722 | 0.696 | 0.709 | 0.571 | 0.782 | 0.660 | 0.759 | 0.777 | 0.768 | 0.757 | 0.686 | | LDA | 1350  1300 | 0.598  0.584 | 0.791  0.809 | 0.681  0.678 | 0.651  0.566 | 0.558  0.475 | 0.516 | 0.482  0.594 | 0.333  0.461 | 0.394  0.519 | 0.522  0.570 | 0.402  0.356 | 0.454  0.438 | 0.581  0.580 | 0.581  0.580 | | 2350 0.671 0.640 0.655 0.524 0.498 0.511 0.456 0.497 0.475 0.584 0.621 0.585 0.585 | | | | | | | | | | | | | | | |   总体上，Joy Anger 恐惧 悲伤  表4：用于检测不同情绪的不同模型的精度量词的比较。表的上部显示了使用（*d，r，ρ*）不同值的NMF技术提出的模型的性能; 分别使用PCA和LDA技术时，中部和底部决定了模型的准确性。  突出显示了每种噪声消除技术的最高*F* 1分数。 |

使用学习技术的情感检测模型的悬念；

* 评估应用于数据集以提取情感的单词袋多层感知器的性能；
* 将多标签预测应用于数据集，并将结果与​​这项工作进行比较，
* 评估文本单元大小（句子，段落，故事）对情感标签准确性的影响；即，将句子分组为较长的单元（例如，段落）并为该较长的单元分配单个标签是否有优势？看来，一个句子本身可能并不总是带有足够的线索来消除其情感的歧义，但其周围的句子可能会提供这种上下文。

# 贡献

我们在叙述性文本中确定了一种高性能的情感识别方法（Kim等，2010年），并仔细地重新实现和表征了该技术，探索了三种不同的消噪或降维技术（NMF，PCA或LDA）的设计空间，探索各种超参数设置。我们的实验表明NMF表现最佳，总*F* 1为0.809。在我们的调查过程中，我们澄清了Kim等人报告的工作的许多实施问题。（2010年），并通过手动添加新术语并使用Wordnet相似关系对WordNet Affect（WNA）（系统中使用的一种语言资源）进行了一些改进。这项工作提出了一些有希望的改进工作的未来方向，包括仔细注解更大的语料库，以及增加WNA或类似的词典来改善情感术语的覆盖范围。我们发布代码和数据以支持将来的工作[[7]](#footnote-7)[4]。[[8]](#footnote-8)

# 致谢

本材料基于美国国土安全部在授予奖号2017-ST-062-000002下的支持工作。本文档中包含的观点和结论是作者的观点和结论，不应解释为必然代表美国国土安全部的明示或暗示的官方政策。

# 参考

Herve Abdi和Lynne J. Williams。2010.′ [Princi pal分量分析。](https://doi.org/10.1002/wics.101)*Wiley跨学科评论：计算统计学*，2：433–459。

Ebba Cecilia Ovesdotter Alm。2008。*对\*文字和语音的影响*。博士 论文，伊利诺伊大学厄本那-香槟分校，伊利诺伊州厄本那-香槟分校。

塞玛·阿曼（Saima Aman）和斯坦·斯帕科维奇（Stan Szpakowicz）。2008年。使用roget同义词库进行细粒度的情感识别。在*第三届国际自然语言处理联合会议论文集：第一卷*，第312-318页，印度海得拉巴。

Anil Bandhakavi，Nirmalie Wiratunga，P Deepak和Stewart Massie。2014年。从#emotional tweets生成词情感词典。在*第三届词汇与计算语义学联合会议论文集（\* SEM 2014）中*，第12-21页，爱尔兰都柏林。

Anil Bandhakavi，Nirmalie Wiratunga，Deepak Padmanabhan和Stewart Massie。2017年[词汇基于特征提取的情感文字classifica重刑。](https://doi.org/10.1016/j.patrec.2016.12.009)*模式识别字母*，93：133–142。

David M Blei，Andrew Y Ng和Michael I Jordan。2003。潜在的狄利克雷分配。*机器学习研究杂志*，3（Jan）：993–1022。

Christos Boutsidis和Efstratios Gallopoulos。2008。基于SVD的初始化：非[负矩阵分解。](https://doi.org/10.1016/j.patcog.2007.09.010)*模式识别*，41（4）：1350–1362。

玛格丽特·布拉德利（Margaret M Bradley）和彼得·朗（Peter J Lang）。1999年。《英语单词情感规范》（ANEW）：使用说明和情感等级。技术报告C-1，佛罗里达州Gainseville的心理生理研究中心。

Fabio Calefato，Filippo Lanubile和Nicole Novielli。2017. [EmoTxt：用于从文本进行情感识别的工具包。](https://doi.org/10.1109/ACIIW.2017.8272591)在*《第七届情感计算与智能国际会议论文集：交互研讨会和演示》（ACIIW 2017）*，第79–80页，德克萨斯州圣安东尼奥。

Rafael A Calvo和Sunghwan Mac Kim。2013。[文字中的情感：维度和分类模型。](https://doi.org/10.1111/j.1467-8640.2012.00456.x)*计算智能*，29（3）：527-543。

埃里克·坎布里亚（Erik Cambria）。2016.[情感计算与情感分析。](https://doi.org/10.1109/MIS.2016.31)*IEEE智能系统*，31（2）：102–107。

埃里克·坎布里亚（Erik Cambria），安德鲁·利文斯通（Andrew Livingstone）和阿米尔·侯赛因（Amir Hussain）。2012。情绪的沙漏。在安娜·埃斯波西托（Anna Esposito）中，安东尼奥·M·埃斯波西托（Antonietta M. Esposito），亚历山德罗·文西亚雷利（Alessandro Vinciarelli），鲁迪格·霍夫曼（Rudiger Hoffmann）和文森特· 穆勒（Vincent M· uller）编着，《*认知行为系统》*，第144-157页。柏林施普林格。作为第7403卷出版，《计算机科学讲义》（LNCS）。

Colin Cherry，Saif M Mohammad和Berry de Bruijn。2012。[二元分类器和潜在序列模型，用于自杀笔记中的情绪检测。](https://doi.org/10.4137/BII.S8933)*生物医学信息学洞察*，5：BII–S8933。

Roddy Cowie，Ellen Douglas-Cowie和Cate Cox。2005。[超越情感原型：使用神经网络进行情感建模的数据库。](https://doi.org/10.1016/j.neunet.2005.03.002)*神经网络*，18（4）：371–388。

查尔斯·达尔文（Charles Darwin）和菲利普·普罗德（Phillip Prodger）。1998年*。*

*人与动物情感的融合*。牛津大学出版社，英国牛津。

保罗·埃克曼。1992a。[有基本情绪吗？](https://doi.org/10.1037/0033-295X.99.3.550)*心理评论*，99（3）：550–553。

保罗·埃克曼。1992b。[关于基本情感的论据。](https://doi.org/10.1080/02699939208411068)*认知与情感*，6（3-4）：169–200。

保罗·埃克曼。1993。[面部表情和情感。](https://doi.org/10.1037/0003-066X.48.4.384)*美国心理学家*，48（4）：384。

克里斯蒂安妮·费尔鲍姆（Christiane Fellbaum），编辑。1998年*。WordNet：一个电子词汇数据库*。麻省理工学院出版社，马萨诸塞州剑桥。

Johnny RJ Fontaine，Klaus R Scherer，Etienne B Roesch和Phoebe C Ellsworth。2007。[情感世界不是二维的。](https://doi.org/10.1111/j.1467-9280.2007.02024.x)*心理科学*，18（12）：1050-1057。

卡洛尔·伊扎德（Carroll E Izard）。2007。基本情绪，自然种类，情绪模式和新范式。*心理科学观点*，2（3）：260–280。

威廉·詹姆斯。1890.*心理学原理*。亨利·霍尔特（Henry Holt）和公司，纽约。

Sunghwan Mac Kim，Alessandro Valitutti和Rafael A Calvo。2010。对文本情感识别的无监督情感模型的评估。在

*NAACL HLT 2010研讨会，有关文本中情感分析和生成的计算方法的会议论文集*，第62–70页，加利福尼亚州洛杉矶。

Matthew W Kreuter，Melanie C Green，Joseph N Cappella，Michael D Slater，Meg E Wise，Doug Storey，Eddie M Clark，Daniel J O'Keefe，Deborah O Erwin，Kathleen Holmes等。2007年。《癌症预防和控制中的叙事交流：指导研究和应用的框架》。*行为医学年鉴*，33（3）：221–235。

George Langroudi，Anna Jourdanous和Ling Li。2018。音乐情感捕获：声音化情感数据。在*社交媒体和在线互动中的情感建模与检测专题讨论会上*，第1-4页，英国利物浦。

李小龙（Daniel D Lee）和塞巴斯蒂安（Sebastian Seung）H。1999年[学习 由非负矩阵factoriza对象的部分和灰。](https://doi.org/10.1038/44565)*自然*，401（6755）：788–791。

雨果·洛夫海姆（Hugo Lovheim）。2012年[一个新的三维模型的情感和单胺neurotransmit TER值。](https://doi.org/10.1016/j.mehy.2011.11.016)*医学假设*，78（2）：341–348。

朱利安·梅拉尔（Julien Mairal），弗朗西斯·巴赫（Francis Bach），让·庞塞（Jean Ponce）和吉列尔莫·萨皮罗（Guillermo Sapiro）。2009。[在线字典学习稀疏编码。](https://doi.org/10.1145/1553374.1553463)在*第26届年度机器学习国际会议论文集中*，第689-696页，加拿大蒙特利尔。

克里斯托弗·D·曼宁（Christopher D Manning），米海·苏迪亚努（Mihai Surdeanu），约翰·鲍尔（John Bauer），

珍妮·罗斯·芬克尔（Jenny Rose Finkel），史蒂芬·贝萨德（Steven Bethard）和戴维·麦克洛斯基（David McClosky）。2014年。斯坦福大学CoreNLP自然语言处理工具包。在*第52辑中*

*计算语言学协会年会（ACL 2014）：系统演示*，第55–60页，美国马里兰州巴尔的摩

DA Medler，A Arnoldussen，JR Binder和MS Seidenberg。2005。威斯康星州感知属性等级（WPAR）数据库。于2020年4月23日从<http://www.neuro.mcw.edu/ratings>检索。

赛义夫·穆罕默德（Saif M Mohammad），斯维特兰娜·基里琴科（Svetlana Kiritchenko）和朱小丹。2013年[NRC-加拿大：建立国家的最先进的鸣叫的情感分析。](https://arxiv.org/abs/1308.6242)*arXiv预印本arXiv：1308.6242*。

赛义夫·穆罕默德（Saif M Mohammad）和彼得·特尼（Peter D Turney）。2010。常见单词和短语引起的情感：使用机械特克创建情感词典。在

*NAACL HLT 2010研讨会，有关文本中情感分析和生成的计算方法的会议论文集*，第26–34页，加利福尼亚州洛杉矶。

基思·奥特利（Keith Oatley）和菲利普·N·约翰逊（Philip N Johnson-Laird）。1987年[为了病房情感的认知理论。](https://doi.org/10.1080/02699938708408362)*认知与情感*，1（1）：29–50。

安德鲁·奥托尼（Andrew Ortony），杰拉尔德·L·克洛（Gerald L Clore）和艾伦·柯林斯（Allan Collins）。

1990.*情绪的认知结构*。剑桥大学出版社，英国剑桥。

Jaak Panksepp，Brian Knutson和Douglas L Pruitt。1998。[迈向情感神经科学。](https://doi.org/10.1007/978-1-4899-1939-7_3)在*情感发展中会发生什么？*，第53-84页。马萨诸塞州波士顿，斯普林格。

W杰罗德·帕罗特（W Gerrod Parrott）。2001年*。《社会心理学中的情感：必读》*。伦敦心理学出版社。

彭浩若，丹尼尔·哈沙比和丹·罗斯。2019.[解决硬性同居问题。](https://arxiv.org/abs/1907.05524)*arXiv预印本arXiv：1907.05524*。

詹姆斯·W·彭纳贝克（James W Pennebaker），玛莎·弗朗西斯（Martha E Francis）和罗杰（Roger J）

展位 2001。语言查询和单词计数（LIWC）软件。

罗伯特·普鲁奇克（Robert Plutchik）。1980.[情绪的一般心理进化论。](https://doi.org/10.1016/B978-0-12-558701-3.50007-7)在*情感理论*编辑罗伯特·普鲁奇克（Robert Plutchik）*中*，第3至33页。Elsevier，阿姆斯特丹，荷兰。

罗伯特·普鲁奇克（Robert Plutchik）。1984年。情感与意象。*心理影像杂志*，8：105–111。

罗伯特·普鲁奇克（Robert Plutchik）。2001年。情绪的本质：人类情绪具有深厚的进化根源，这一事实可以解释其复杂性并为临床实践提供工具。*美国科学家*，89（4）：344–350。

詹姆斯·罗素（James A Russell）。1980. [af-的外接模型](https://doi.org/10.1037/h0077714)

[效果。](https://doi.org/10.1037/h0077714)*人格与社会心理学杂志*，39（6）：1161。

克劳斯·谢勒（Klaus R Scherer）。2005。[什么是情感？以及如何测量它们？](https://doi.org/10.1177/0539018405058216)*社会科学信息*，44（4）：695–729。

沙迪·沙欣（Shadi Shaheen），瓦西姆·哈吉（Wassim El-Hajj），哈兹姆·哈吉（Hazem Hajj）和谢迪·艾尔巴索尼（Shady Elbassuoni）。2014。基于自动生成的规则从文本进行情感识别。在*2014 IEEE国际数据挖掘会议研讨会论文集*，第383–392页，中国深圳。

菲利普·谢弗（Phillip Shaver），朱迪思·施瓦茨（Judith Schwartz），唐纳德·基森（Donald Kirson）和卡里·奥康纳（Cary O'connor）。1987年。[情感知识：原型方法的进一步探索。](https://doi.org/10.1037/0022-3514.52.6.1061)*人格与社会心理学杂志*，52（6）：1061-1086。

Carlo Strapparava和Rada Mihalcea。2008年。[学习识别文本中的情绪。](https://doi.org/10.1145/1363686.1364052)在*2008 ACM应用计算研讨会论文集*，第1556-1560页，巴西塞阿拉省福塔莱萨。

Carlo Strapparava和Alessandro Valitutti。2004年。

WordNet影响：Wordnet的情感扩展。在*第四届国际语言资源与评估国际会议论文集（LREC'04）中*，第1083–1086页，葡萄牙里斯本。

Jared Suttles和Nancy Ide。2013。[对离散分类二进制值进行情感分类的远程监督。](https://doi.org/10.1007/978-3-642-37256-8_11)在*国际智能文本处理和计算语言学国际会议论文集*，第121–136页，德国柏林。

智腾，富士人和黑岩真吾。2007。基于粗糙集理论和支持向量机的文本情感识别。在*2007年国际自然语言处理和知识工程国际会议论文集中*，第36–41页，北京，中国。

德寿凉子，犬井健太郎和松本雄司。2008年。使用从网络中提取的大量示例进行情感分类。在*第22届国际计算语言学国际会议论文集：第1卷*，第881–888页，英国曼彻斯特。

Ozlem Uzuner，Andreea Bodnari，沉淑英，Tyler Forbush，John Pestian和Brett R South。2012年。[评估电子病历共参考解决方案的最新状态。](https://doi.org/10.1136/amiajnl-2011-000784)*杂志*

*美国医学信息学协会*，19（5）：786–791。

张小，李文中，陈旭和卢桑鲁。2018. [Moodexplorer：通过智能手机感应实现复合情感检测。](https://doi.org/10.1145/3161414)*ACM关于交互，移动，可穿戴和无处不在技术的论文集*，1（4）：1–30。

1. 原文

提供更好的翻译建议

1. 可以从[https：// doi](https://doi.org/10.34703/gzx1-9v95/03RERQ)下载代码和数据[。](https://doi.org/10.34703/gzx1-9v95/03RERQ)

   [org / 10.34703 / gzx1-9v95 / 03RERQ](https://doi.org/10.34703/gzx1-9v95/03RERQ) [↑](#footnote-ref-1)
2. [↑](#footnote-ref-2)
3. [http://www.affective-sciences.org/ researchmaterial](http://www.affective-sciences.org/researchmaterial) [↑](#footnote-ref-3)
4. [↑](#footnote-ref-4)
5. Kim等。（2010年）报告了1,093个句子，但我们发现并删除了数据中重复的三个句子。 [↑](#footnote-ref-5)
6. [↑](#footnote-ref-6)
7. 代码和数据可以从[https：// doi](https://doi.org/10.34703/gzx1-9v95/03RERQ)下载[。](https://doi.org/10.34703/gzx1-9v95/03RERQ)

   [org / 10.34703 / gzx1-9v95 / 03RERQ](https://doi.org/10.34703/gzx1-9v95/03RERQ) [↑](#footnote-ref-7)
8. [↑](#footnote-ref-8)