

20. 情感、感情和内涵的词汇

Some day we'll be able to measure the power of words

Maya Angelou

在本章中，我们将转向解释**感情含义(affective meaning)**的工具，扩展我们在第四章中对情感分析的研究。我们使用**感情(affective)**这个词，遵循了**感情计算(affective computing)**的传统(Picard, 1995)来表示情绪、情感、性格、语气和态度。与感情含义密切相关的是主观性：对说话者或作者的评价、意见、情绪和推测的研究(Wiebe 等人, 1999)。

感情含义应该如何定义?一种有影响力的感情状态类型来自 Scherer(2000)，他用认知实现和时间进程等因素来定义每一类感情状态(图 20.1)。

- 情绪(Emotion)**: 对具有重大意义的外部或内部事件的评价所产生的相对短暂的反应。
(愤怒、悲伤、喜悦、恐惧、羞愧、骄傲、得意、绝望)
- 语气(Mood)**: 弥漫的感情状态，最显著的是主观感觉的变化，强度低但持续时间相对较长，通常没有明显的原因。
(愉快的、阴郁的、易怒的、无神的、沮丧的、愉快的)
- 人际立场(Interpersonal stance)**: 在特定的互动中对另一个人所采取的感情立场，使得该情境中的人际交流充满色彩。
(遥远的、冷淡的、温暖的、支持的、轻蔑的、友好的)
- 态度(Attitude)**: 对物体或人的相对持久的、有感情色彩的信念、偏好和成见。
(喜欢、爱、恨、重视、渴望)
- 人格特质(Personality traits)**: 一个人典型的情绪丰满、稳定的人格气质和行为倾向。
(紧张的、焦虑的、鲁莽的、忧郁的、敌对的、嫉妒的)

图 20-1: 感情(affective)状态的 Scherer 类型

我们可为每一种感情状态设计提取器。第四章已经介绍了**情感分析(sentiment analysis)**，即提取作者在文本中表达的积极或消极倾向。这与 Scherer 的类型学的**态度(attitudes)**提取相一致:从消费者对书籍或电影的评论、报纸社论或博客或推特上的公众情感等感情丰富的文本中找出人们喜欢什么或不喜欢什么。

情绪(emotion)和**语气(moods)**的检测在如下三方面是有用的：在与教学系统互动时，学生是否感到困惑、沉迷或确定；求助热线的呼叫者是否感到沮丧；某人的博客文章或推文是否表示泄气。例如，在小说中发现恐惧这样的情绪，可以帮助我们追踪恐惧的群体或情境，以及随着时间的推移恐惧如何变化。

在从人与人之间的对话中提取信息时，检测不同的**人际立场(interpersonal stances)**是很有用的。这里的目标是在采访或友好的对话中发现友好或尴尬的立场，例如总结会议或找到人们特别兴奋或参与的对话部分，对话热点，可以帮助会议总结。检测用户的个性——比如用户是否外向，或者他们对体验的开放程度——可以帮助改进会话施事者，如果它们符合用户的个性期望，似乎会工作得更好(Mairesse 和 Walker, 2008)。感情是生成和识别的重要因素,对感情进行综合对于不同领域的会话施事者来说很重要，包括儿童故事书或电脑游戏等教育辅导。

第四章介绍了朴素贝叶斯分类方法在文档情感分类中的应用。各种分类器已经成功地应用于许多这样的任务，使用训练集中的所有单词作为分类器的输入，然后该分类器确定文本的感情状态。

在本章中，我们着眼于替代模型，在该模型中，我们不是将每个单词都用作特征，而是仅关注某些单词，这些单词带有特别强烈的感情或情感暗示。我们称这些单词为感情词汇或情感词汇。这些词汇以关于语义的事实为前提：单词具有感情含义或**内涵(connotation)**。内涵在不同领域具有不同的含义，但是在这里，我们用它来表示单词含义的某些方面，这些方面与作者或读者的情绪，情感，观点或评价有关。内

涵词汇除了可以帮助确定文本的感情状态外，还可以用于其他类型的感情任务以及计算机社会科学分析。

在接下来的章节中，我们将介绍情绪的基本理论，说明情感词汇是情绪词汇的一种特殊情况，并提出一些有用的词汇。然后，我们调查了三种构建词汇表的方法：人工标签、半监督和监督。最后，我们转向其他类型的感情含义，如人格、立场和以实体为中心的感情，并引入内涵框架。

20.1. 定义情绪

情绪是最重要的感情类别之一，Scherer(2000)将**情绪(emotion)**定义为“对具有重要意义的外部或内部事件评价时的相对短暂的反应”。

检测情绪有可能改善许多语言处理任务。自动检测评论或客户反应中的情绪(愤怒、不满、信任)可以帮助企业认识到特定的问题领域或进展顺利的领域。情感识别可以帮助对话系统，如辅导系统检测到一个学生不高兴、无聊、犹豫、自信，等等。情绪可以在医学信息任务中发挥作用，如检测抑郁或自杀意图。检测小说中人物所表达的情绪，可能对理解社会在不同时期对不同社会群体的看法起着重要作用。

有两个广为接受的情绪理论家族。我们将为两种理论介绍词汇。在一个家族中，情绪被视为数量有限的固定原子单位，并且从中产生其他情感，通常被称为**基本情绪(basic emotions)**(Tomkins 1962, Plutchik 1962)。在这一系列理论中，最著名的也许是 Ekman 提出的 6 种情绪(例如，见 Ekman1999)，这是可能在所有文化中普遍存在的一组情绪：惊奇、幸福、愤怒、恐惧、厌恶、悲伤。另一个原子理论是“情绪轮”(Plutchik, 1980 年)，由四个相反的偶对组成的 8 种基本情绪：喜悦-悲伤、愤怒-恐惧、信任-厌恶、期待-惊奇，以及从中得出的情绪，如图 20.2 所示：

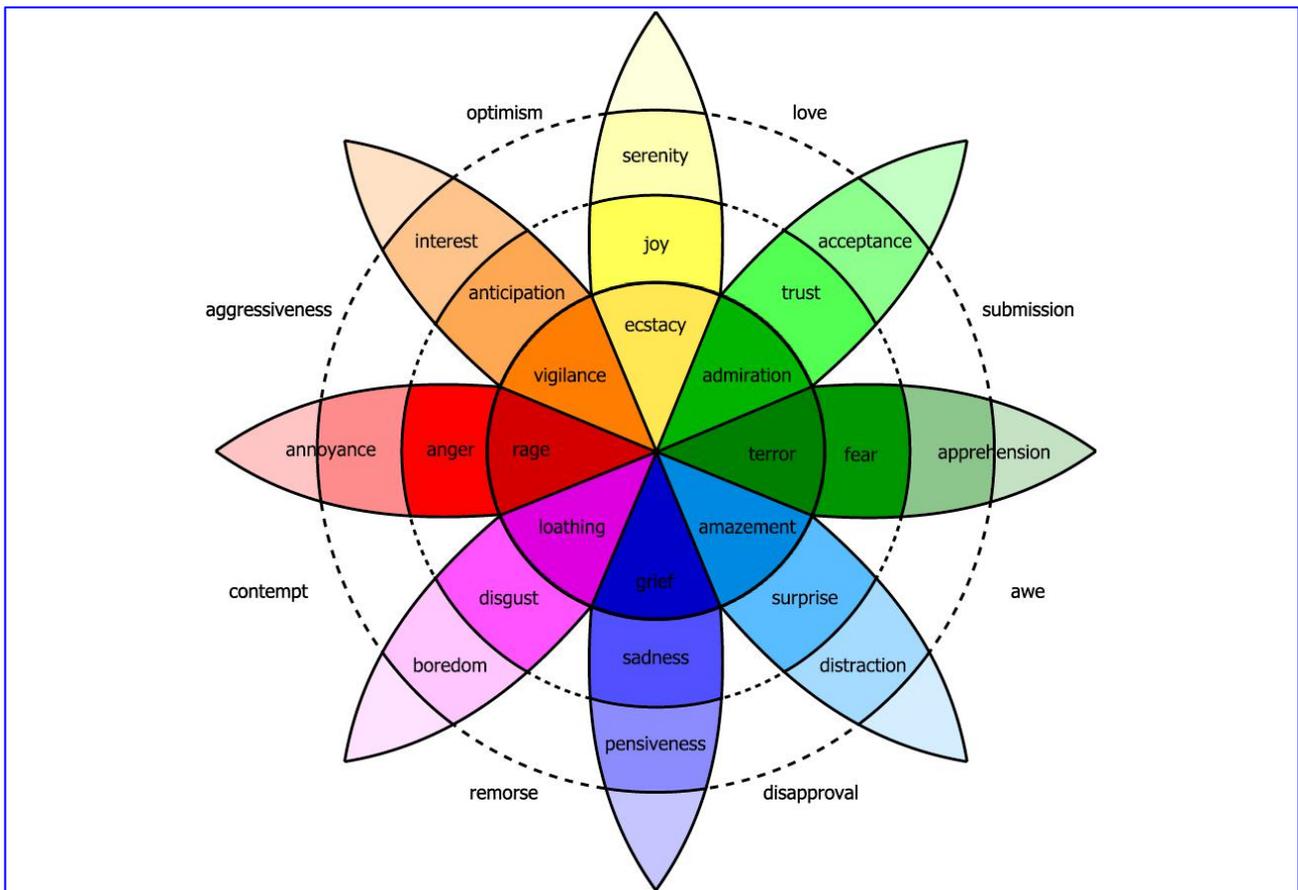


图 20-2: Plutchik 情绪轮-8 种基本情绪

第二类情绪理论将情感视为 2 个或 3 个维度的空间(Russell, 1980)。大多数模型包括**效价(valence)**和**唤起(arousal)**这两个维度，许多模型还增加了第三个维度，**支配(dominance)**。这些可以定义为：

效价(V): 刺激的愉悦感

唤起(A): 刺激所激起的情绪强度

支配(D): 刺激所施加的控制程度

情感可以被看作是第二种情绪的空间点的特殊情况。尤其是效价维度，衡量一个词有多令人愉快或不愉快，经常直接被用来衡量情感。

20.2. 可用的情感和感情词汇

各种各样的**感情(affect)**词汇已经被创造和发布。最基本的词汇根据语义变化的一个维度来标注单词，通常称为“**情感(sentiment)**”或“**效价(valence)**”。

在最简单的词汇中，此维度以二元方式表示，其中有一个用于正面的单词表和一个用于负面的单词表。最古老的是**普通调查员(General Inquirer, GI)**(Stone 等人, 1966)，它是根据内容分析和对单词意义的认知心理学的早期研究而获得的(Osgood 等人, 1957)。General Inquirer 有一个包含 1915 个正面词的词汇和一个包含 2291 个负面词的词汇(以及下面讨论的其他词汇)。MPQA 主观词汇(Wilson 等, 2005)具有从先前词汇中抽取的 2718 个正面词和 4912 个负面词，以及一系列主观单词和短语的引导表(Riloff 和 Wiebe, 2003 年)。词汇中的每个词条都经过手工标签以表达情感，还标签有可靠性(强烈主观或弱主观)。Hu 和 Liu(2004b)的极性(polarity)词汇提供了 2006 正面词和 4783 负面词，这些词来自产品评论，并使用 WordNet 的引导方法进行了标签，见图 20.3。

Positive	admire, amazing, assure, celebration, charm, eager, enthusiastic, excellent, fancy, fantastic, frolic, graceful, happy, joy, luck, majesty, mercy, nice, patience, perfect, proud, rejoice, relief, respect, satisfactorily, sensational, super, terrific, thank, vivid, wise, wonderful, zest
Negative	abominable, anger, anxious, bad, catastrophe, cheap, complaint, condescending, deceit, defective, disappointment, embarrass, fake, fear, filthy, fool, guilt, hate, idiot, inflict, lazy, miserable, mourn, nervous, objection, pest, plot, reject, scream, silly, terrible, unfriendly, vile, wicked

图 20-3: 普通调查员、主观词汇、极性词汇

图注: General Inquirer(Stone 等人, 1966)、MPQA 主观词汇(Wilson 等人, 2005)和 Hu and Liu(2004b)的极性词汇中的一些情感一致的单词。

在所有三个感情维度上，为每个单词分配一个值的词汇要比这些情感词汇稍微普遍一些。NRC 的效价、唤起和支配(VAD)词汇(Mohammad, 2018a)将效价，唤起和支配的分数分配给 20,000 个单词。图 20.4 中显示了一些示例。

Valence		Arousal		Dominance	
vacation	.840	enraged	.962	powerful	.991
delightful	.918	party	.840	authority	.935
whistle	.653	organized	.337	saxophone	.482
consolation	.408	effortless	.120	discouraged	.009
torture	.115	napping	.046	weak	.045

图 20-4: 样本词在 Mohammad 情绪维度上的值

NRC 单词情绪关联词汇，也称为 **EmoLex**(Mohammad 和 Turney, 2013)，使用了上面定义的 Plutchik(1980)的 8 种基本情绪。该词汇包括大约 14,000 个单词，包括以前词汇中的单词以及常见的名词、动词、副词和形容词。词汇中一些示例单词的值：

Word	anger	anticipation	disgust	fear	joy	sadness	surprise	trust	positive	negative
reward	0	1	0	0	1	0	1	1	1	0
worry	0	1	0	1	0	1	0	0	0	1
tenderness	0	0	0	0	1	0	0	0	1	0
sweetheart	0	1	0	0	1	1	0	1	1	0
suddenly	0	0	0	0	0	0	1	0	0	0
thirst	0	1	0	0	0	1	1	0	0	0
garbage	0	0	1	0	0	0	0	0	0	1

对于一个较小的 5814 个单词的集合, NRC 情绪/感情强度词汇表(Mohammad, 2018b)包含了愤怒、恐惧、快乐和悲伤的关联的真实值分数;图 20.5 给出了示例。

Anger		Fear		Joy		Sadness	
outraged	0.964	horror	0.923	superb	0.864	sad	0.844
violence	0.742	anguish	0.703	cheered	0.773	guilt	0.750
coup	0.578	pestilence	0.625	rainbow	0.531	unkind	0.547
oust	0.484	stressed	0.531	gesture	0.387	difficulties	0.421
suspicious	0.484	failing	0.531	warms	0.391	beggar	0.422
nurture	0.059	confident	0.094	hardship	.031	sing	0.017

图 20-5: 对愤怒、恐惧、快乐和悲伤词汇的情绪强度样本

语言查询和单词计数(Linguistic Inquiry and Word Count, LIWC), 是一个被广泛使用的包含超过 2300 个单词的 73 个词汇集(Pennebaker 等人, 2007), 旨在捕捉与社会心理任务相关的词汇意义方面。除了与情感相关的词汇, 如消极情绪(坏、奇怪、恨、问题、强硬)和积极情绪(爱、美好、甜蜜), LIWC 还包括了愤怒、悲伤、认知机制、知觉、试探和抑制等类别的词汇, 如图 20.6 所示。

Positive Emotion	Negative Emotion	Insight	Inhibition	Family	Negate
appreciat*	anger*	aware*	avoid*	brother*	aren't
comfort*	bore*	believe	careful*	cousin*	cannot
great	cry	decid*	hesitat*	daughter*	didn't
happy	despair*	feel	limit*	family	neither
interest	fail*	figur*	oppos*	father*	never
joy*	fear	know	prevent*	grandf*	no
perfect*	griev*	knew	reluctan*	grandm*	nobod*
please*	hate*	means	safe*	husband	none
safe*	panic*	notice*	stop	mom	nor
terrific	suffers	recogni*	stubborn*	mother	nothing
value	terrify	sense	wait	niece*	nowhere
wow*	violent*	think	wary	wife	without

图 20-6: LIWC 中 73 个词汇类别中的 5 个样本

图注: 其中*表示前面的字母是一个单词前缀, 所有带有该前缀的单词都包含在类别中(Pennebaker 等人, 2007)。

还有其他各种各样的手工构建的感情词汇。一般查询包括诸如强与弱、主动与被动、夸大与低估等维度的额外词汇, 以及诸如快乐、痛苦、美德、邪恶、动机和认知取向等类别的词汇。

对各种任务来说, 另一个有用的特征是**具体(concrete)**单词(如 banana 或 bathrobe)与**抽象(abstract)**单词(如 belief 和 although)之间的区别。Brysaert 等人(2014)的词汇使用众包的方法对 4 万个单词的具象性进行了从 1 到 5 的评分, 例如给 banana、bathrobe 和 bagel 打 5 分, belief 打 1.19 分, although 打 1.07 分, 以及中间的单词, 诸如 brisk, 打 2.5 分。

20.3. 通过人工标签创建感情词汇

最早用来建立感情词汇的方法是让人给每个词贴上标签, 这一方法至今仍在使用。现在最普遍的做法是通过**众包(crowdsourcing)**来完成:把任务分成小块, 然后分发给大量的评注者。让我们来看看这两个来自大众的情绪词汇在方法论上的选择。

NRC 情绪词典(EmoLex) (Mohammad 和 Turney, 2013), 用两个步骤标记情绪。为了确保评注者判断的是这个词的正确意思, 他们首先回答了一个同义词的多项选择题, 这个问题引发了这个词的正确意义(不需要评注者阅读可能会造成混淆的意义定义)。这些是使用以下两类中心词自动创建的: 与 Macquarie 字典中所讨论的意义的同义词库类别相关的中心词; 3 个随机干扰项类别的中心词。一个例子:

Which word is closest in meaning (most related) to **startle**?

- automobile • shake • honesty • entertain

对于每一个词(例如 **startle**), 评注者被要求对这个词与 8 种情绪(joy, fear, anger 等)之间的联系进行评级。这些关联在 not, weakly, moderately, 和 strongly 的尺度(scale)上被评级。除去异常值评级(Outlier ratings), 然后为每个术语分配大多数评注者选择的类别, 并通过选择更强的强度来打破纽带关系, 然后将 4 个级别映射到每个单词的二元标签中(否和弱映射到 0, 中度和强度对应于 1)。

NRC 的 VAD 词汇(Mohammad, 2018a)是通过从先前的词汇中选择单词和情绪符号, 并用**最好-最差缩放(best-worst scaling)**的群体外包对它们进行评注而构建的(Louviere 等人 2015, Kiritchenko 和 Mohammad 2017)。在最好-最差缩放中, 给评注者 N 项(通常是 4 项), 并被问及就某些属性而言, 哪项是最好的(最高的), 哪项是最差的(最低的)。用来描述尺度(scales)末端的一组词取自先前的文献。例如, 对于效价, 评级者(raters)被要求:

Q1. Which of the four words below is associated with the

MOST happiness / pleasure / positiveness / satisfaction / contentedness / hopefulness

OR LEAST unhappiness / annoyance / negativeness / dissatisfaction / melancholy / despair?

(Four words listed as options.)

Q2. Which of the four words below is associated with the

LEAST happiness / pleasure / positiveness / satisfaction / contentedness / hopefulness

OR MOST unhappiness / annoyance / negativeness / dissatisfaction / melancholy / despair?

(Four words listed as options.)

词汇表中每个单词的得分是该项被选为最好(最高的 V/A/D)的次数的比例减去该项被选为最差(最低的 V/A/D)的次数的比例。评注之间的一致性通过**对半可靠性(split-half reliability)**进行评估: 将语料库对半分并计算两半之间评注之间的相关性。

20.4. 感情词汇的半监督归纳

学习情感词汇的另一种常见方法是从一组定义语义轴两极的种子词(如 good 或 bad)开始, 然后通过这两种种子集的相似性来标签每个单词 w 。在这里我们总结了两族基于种子的半监督词汇归纳算法: 基于轴的和基于图的算法。

20.4.1. 语义轴方法

最著名的词汇归纳方法之一, Turney 和 Littman(2003)算法, 是给定种子词, 例如 good 或 bad, 然后为每个单词 w 贴标签, 来衡量它既是如何与 good 相似的, 又是如何与 bad 不同的。在这里, 我们描述了基于计算语义轴的算法的略微扩展, 这归功于 An 等人(2018)。

第一步, 我们手工选择种子词。有两种方法可以解决单词在不同上下文中的感情不同的事实:(1)从一个大型种子词汇开始, 依靠归纳算法对其进行微调到领域(domain), 或者(2)为不同的体裁选择不同的种子词汇。Hellrich 等人(2019)认为, 要对跨越不同历史时期的感情建模, 从一个大型的现代感情字典开始比调整为跨时间稳定的小种子集更好。作为第二种方法的一个例子, Hamilton 等人(2016a)定义一组用于一般情感分析的种子词, 一组用于 Twitter 的种子词, 还有一组用于金融文本的种子词:

Domain	Positive seeds	Negative seeds
General	good, lovely, excellent, fortunate, pleasant, delightful, perfect, loved, love, happy	bad, horrible, poor, unfortunate, unpleasant, disgusting, evil, hated, hate, unhappy
Twitter	love, loved, loves, awesome, nice, amazing, best, fantastic, correct, happy	hate, hated, hates, terrible, nasty, awful, worst, horrible, wrong, sad
Finance	successful, excellent, profit, beneficial, improving, improved, success, gains, positive	negligent, loss, volatile, wrong, losses, damages, bad, litigation, failure, down, negative

第二步, 我们计算每个极点词的嵌入。这些嵌入可以是现成的 word2vec 嵌入, 或者可以在特定语料库上直接计算(例如, 如果目标是金融词汇, 则使用金融语料库), 或者我们可以微调现成的嵌入到语料库。如果我们有一个非常特定的文本体裁, 但没有足够的数据来训练良好的嵌入, 微调就特别重要。在微调方

面, 我们从现成的嵌入式开始, 如 `word2vec`, 并继续在小型目标语料库上训练它们。

一旦我们有了每个极点词的嵌入, 我们通过取每个种子词嵌入的质心(centroid)来创建一个代表每个极点的嵌入; 回想一下, 质心是均值的多维版本。给定一组正面种子词 $S^+ = \{E(w_1^+), E(w_2^+), \dots, E(w_n^+)\}$, 负面种子词 $S^- = \{E(w_1^-), E(w_2^-), \dots, E(w_m^-)\}$, 极点质心为:

$$\begin{aligned} \mathbf{V}^+ &= \frac{1}{n} \sum_{i=1}^n E(w_i^+) \\ \mathbf{V}^- &= \frac{1}{m} \sum_{i=1}^m E(w_i^-) \end{aligned} \quad (20.1)$$

由极点定义的语义轴的计算只需减去两个向量:

$$\mathbf{V}_{\text{axis}} = \mathbf{V}^+ - \mathbf{V}^- \quad (20.2)$$

\mathbf{V}_{axis} , 即语义轴, 是正向情感的矢量。最后, 我们通过余弦相似度计算出正向情感方向与 \mathbf{w} 嵌入方向的夹角。更高的余弦意味着 \mathbf{w} 更倾向于 S^+ 而不是 S^- 。

$$\begin{aligned} \text{score}(w) &= (\cos(E(w), \mathbf{V}_{\text{axis}})) \\ &= \frac{E(w) \cdot \mathbf{V}_{\text{axis}}}{\|E(w)\| \|\mathbf{V}_{\text{axis}}\|} \end{aligned} \quad (20.3)$$

如果一本带有情感评分的单词字典就足够了, 那么我们就完成了! 或者, 如果我们需要将词汇分为正面词汇和负面词汇, 那么我们可以使用一个阈值或其他方法来给予我们离散词汇。

20.4.2. 标签传播

另一种方法是通过在图上传播情感标签来定义词汇, 这是 Hatzivassiloglou 和 McKeown(1997) 早期工作中提出的想法。我们将描述 Hamilton 等人(2016a) 的简单情感传播(SentProp)算法, 该算法有四个步骤:

1. 定义一个图: 给定词的嵌入, 通过连接每个词与其 k 个最近的邻居(根据余弦相似度)建立一个加权词图。设单词 w_i 和 w_j 之间的边的权值为:

$$\mathbf{E}_{i,j} = \arccos \left(-\frac{\mathbf{w}_i^\top \mathbf{w}_j}{\|\mathbf{w}_i\| \|\mathbf{w}_j\|} \right). \quad (20.4)$$

2. 定义一个种子集: 选择正面种子词和负面种子词。

3. 从种子集中传播极性: 现在我们在这个图上执行随机漫步, 从种子集中开始。在随机漫游中, 我们从一个节点开始, 然后选择与边概率成比例的概率移动到一个节点。一个单词的极性分数与从种子集到该单词的随机漫步概率成正比(图 20.7)。

4. 创建单词分数: 我们从正面和负面种子集走来, 得到结果是正($\text{rawscore}^+(w_i)$)和负($\text{rawscore}^-(w_i)$)原始标签分数。然后我们将这些值组合成正极性分数如下:

$$\text{score}^+(w_i) = \frac{\text{rawscore}^+(w_i)}{\text{rawscore}^+(w_i) + \text{rawscore}^-(w_i)} \quad (20.5)$$

将分数标准化, 使其在一个语料库中均值和单位方差为零通常是有帮助的。

5. 给每个分数赋值信度: 因为情绪分数受到种子集的影响, 我们想知道如果使用不同的种子集, 单词的分数会发生多大的变化。我们可以使用自举抽样来获得置信区域, 通过计算在正面种子集和负面种子集的随机子集上的传播次数 B (例如使用 $B=50$, 每次从 10 个种子词中选择 7 个)。自举取样极性分数的标准偏差给出了一个置信度量。

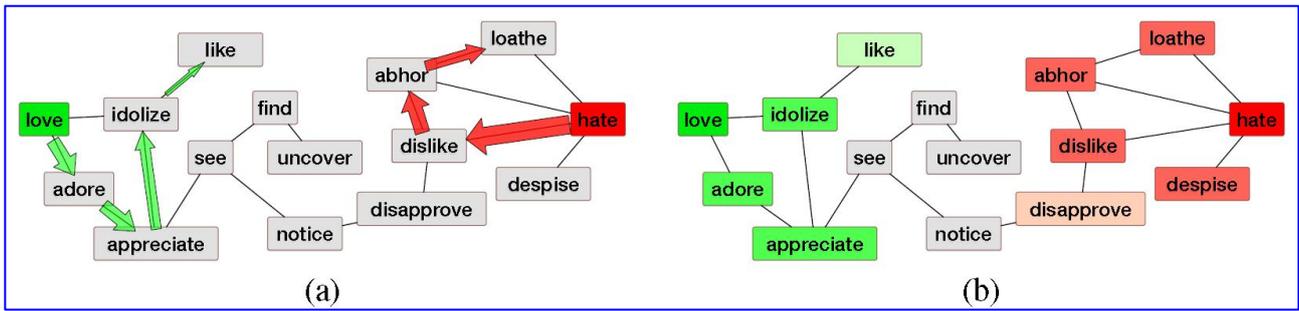


图 20-7: SentProp 算法的直观感受

图注: 图(a)从种子词中随机漫步。图(b)根据随机行走的次数分配极性分数(在这里用绿色或红色表示)。

20.4.3. 其它方法

半监督算法的核心是与种子词相似度的度量。Turney 和 Littman(2003)和 Hamilton 等人(2016a)在上述方法中使用了嵌入余弦作为距离度量:如果单词的嵌入具有带有正面种子的高余弦和带有负面种子的低余弦,则基本上将其标记为正余弦。其他方法除了嵌入余弦外,还选择了其他类型的距离度量。

例如, Hatzivassiloglou 和 McKeown(1997)算法使用了语法线索;如果两个形容词经常由 **and** 连接,很少由 **but** 连接,那么它们被认为是相似的。这是基于一种直觉,即由单词连接而成的形容词往往具有相同的极性;正面形容词通常与正面并列,负面形容词通常与负面并列:

fair and legitimate, corrupt and brutal (公平而合法, 腐败且残忍)

但与负面连用的正面形容词较少:

***fair and brutal, *corrupt and legitimate**

相比之下,由 **but** 连接的形容词可能具有相反的极性:

fair but brutal

相反极性的另一个线索来自形态否定(**un-**, **im-**, **-less**)。词根相同但在形态上否定的不同的形容词(**adequate/adequate, thoughtful/thoughtless**)往往具有相反的极性。另一种寻找与种子词具有相似极性的词的方法是使用 WordNet 这样的同义词库(Kim 和 Hovy 2004, Hu 和 Liu 2004b)。一个单词的同义词可能有相同的极性,而一个单词的反义词可能有相反的极性。在构建种子词汇表之后,每个词汇表将按如下方式更新,可能是迭代的。

Lex⁺: 增加肯定词的同义词(**well**)和否定词的反义词(如 **fine**)

Lex⁻: 增加否定词的同义词(**awful**)和肯定词的反义词(如 **evil**)

该算法的一个扩展将极性分配到 WordNet 的含义,称为 **SentiWordNet**(Baccianella 等人, 2010)。

图 20.8 显示了一些例子。

Synset	Pos	Neg	Obj
good#6 'agreeable or pleasing'	1	0	0
respectable#2 honorable#4 good#4 estimable#2 'deserving of esteem'	0.75	0	0.25
estimable#3 computable#1 'may be computed or estimated'	0	0	1
sting#1 burn#4 bite#2 'cause a sharp or stinging pain'	0	0.875	0.125
acute#6 'of critical importance and consequence'	0.625	0.125	0.250
acute#4 'of an angle; less than 90 degrees'	0	0	1
acute#1 'having or experiencing a rapid onset and short but severe course'	0	0.5	0.5

图 20-8: 来自 SentiWordNet 3.0 的示例

图注: 注意同形词的区别:estimable#3 是纯客观的,而 estimable#2 是正面的;acute 可以是正面(acute#6)、负面(acute#1)或中性(acute#4)。

在该算法中,极性被分配给整个同义词库而不是单词。从与 7 个正面词相关的所有同义词库构建一个正面词库,从与 7 个负面词相关的同义词库构建一个负面词库。然后根据这些数据训练分类器,以获取 WordNet 的注释,并确定所定义的含义是正面的,负面的还是中立的。进一步的步骤(涉及随机漫游算法)

为每个 WordNet 同义词库的正面、负面和中立程度分配分数。

总之，半监督算法，针对维度的两极，使用人类定义的种子词集，并使用类似度量，如嵌入余弦、协调、形态学或词典结构，通过它们与正面种子的相似程度和与负面种子的不同程度来为单词评分。

20.5. 单词情感的监督学习

半监督方法只需要最少的人工监督(以种子组的形式)。但有时世界上存在着一种可以利用的监督信号。其中一个信号就是与在线评论相关的分数。

web 包含了大量关于餐馆、电影、书籍或其他产品的在线评论，每个评论都有评论的文本以及相关的评论评分:评分范围从 1 到 5 星，或 1 到 10 星。图 20.9 显示了从餐厅、书籍和电影评论中提取的样本。

我们可以将这一评分作为监督：正面词汇更有可能出现在 5 星评论中；负面词汇在 1 星评论中。这种监督方式允许我们给一个单词分配一个更复杂的极性（而不仅仅是一种二元极性）：它在星(或其他分数)上的分布。因此，在一个 10 星系统中，可以用 10 元组来表示每个单词的情绪，每个数字代表单词与极性水平的关联。这种关联可以是原始计数，或可能的 $P(w|c)$ ，或计数的其他函数，对于每个 c 类从 1 到 10。

例如，在 IMDb 数据集之中，我们可以计算单词 `disappoint(ed/ing)` 出现 1 星评价的似然度：单词 `disappoint(ed/ing)` 1 星评价出现的次数(8,557)，除以 1 星评价出现的单词总数(25,395,214)，因此 $P(w=\text{disappointing} | c=1)$ 的 IMDb 估计为 $8557/25395214=0.0003$ 。对这一权重(归一化似然度)稍加修改，就可以用作一种启发性的可视化(Potts, 2011)¹：

$$P(w|c) = \frac{\text{count}(w, c)}{\sum_{w \in C} \text{count}(w, c)}$$

$$\text{PottsScore}(w) = \frac{P(w|c)}{\sum_c P(w|c)} \quad (20.6)$$

将 0.0003 的 IMDb 估计值 $P(\text{disappointing} | 1)$ ，除以所有类别上的似然度 $P(w | c)$ 总和，给出的 Potts 得分为 0.10。因此，单词 `disappointing` 与向量[.10, .12, .14, .14, .13, .11, .08, .06, .06, .05]相关联。Potts 图(Potts, 2011 年)是这些单词得分的可视化结果，代表了 Potts 图的单词先验情感，作为评级类别的分布。

Movie review excerpts (IMDb)	
10	A great movie. This film is just a wonderful experience. It's surreal, zany, witty and slapstick all at the same time. And terrific performances too.
1	This was probably the worst movie I have ever seen. The story went nowhere even though they could have done some interesting stuff with it.
Restaurant review excerpts (Yelp)	
5	The service was impeccable. The food was cooked and seasoned perfectly... The watermelon was perfectly square ... The grilled octopus was ... mouthwatering...
2	...it took a while to get our waters, we got our entree before our starter, and we never received silverware or napkins until we requested them...
Book review excerpts (GoodReads)	
1	I am going to try and stop being deceived by eye-catching titles. I so wanted to like this book and was so disappointed by it.
5	This book is hilarious. I would recommend it to anyone looking for a satirical read with a romantic twist and a narrator that keeps butting in.
Product review excerpts (Amazon)	
5	The lid on this blender though is probably what I like the best about it... enables you to pour into something without even taking the lid off! ... the perfect pitcher! ... works fantastic.
1	I hate this blender... It is nearly impossible to get frozen fruit and ice to turn into a smoothie... You have to add a TON of liquid. I also wish it had a spout ...

图 20-9: 摘自不同评论网站的一些评论

图注：除了 IMDb 的评分为 1 - 10 星，其他评分均为 1 - 5 星。

¹ Potts 表明，如果我们做出错误但简化的假设，即所有类别 c 具有相同的概率，则归一化似然度是后验 $P(c|w)$ 的估计。

图 20.10 显示了 3 个正标量和 3 个负标量形容词的 Potts 图。注意，强正标量的曲线具有字母 J 的形状，而强负标量的曲线看起来像是反向 J。相反，弱正标量和负标量具有驼峰形状，最大值低于均值(弱负值词诸如 *disappointing*)或高于平均数(弱正值词诸如 *good*)。这些形状提供了感情含义的启发性类型。

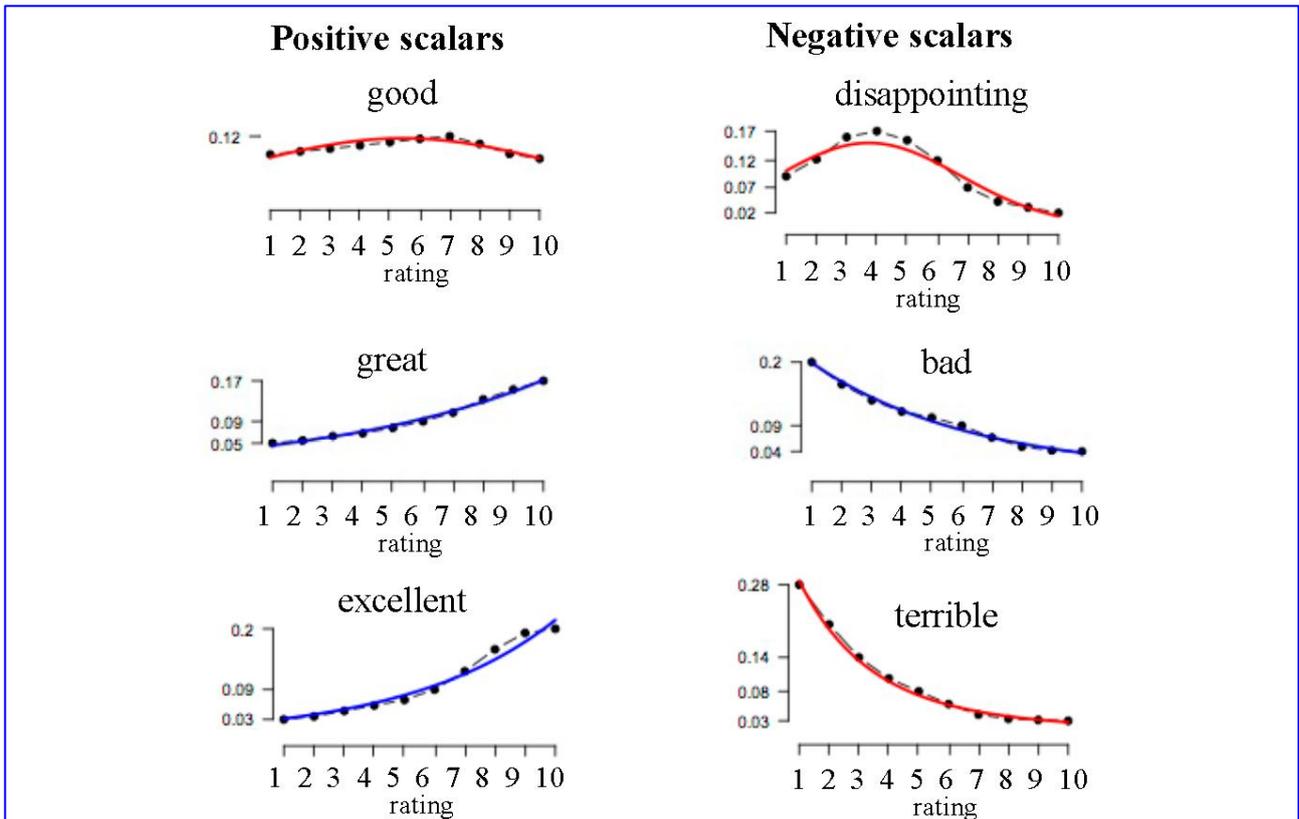


图 20-10: 正面、负面标量形容词的 Potts 图

图注: 显示了强正面、负面形容词的 J 形和反 J 形, 以及较弱极化形容词的峰形(Potts, 2011)。

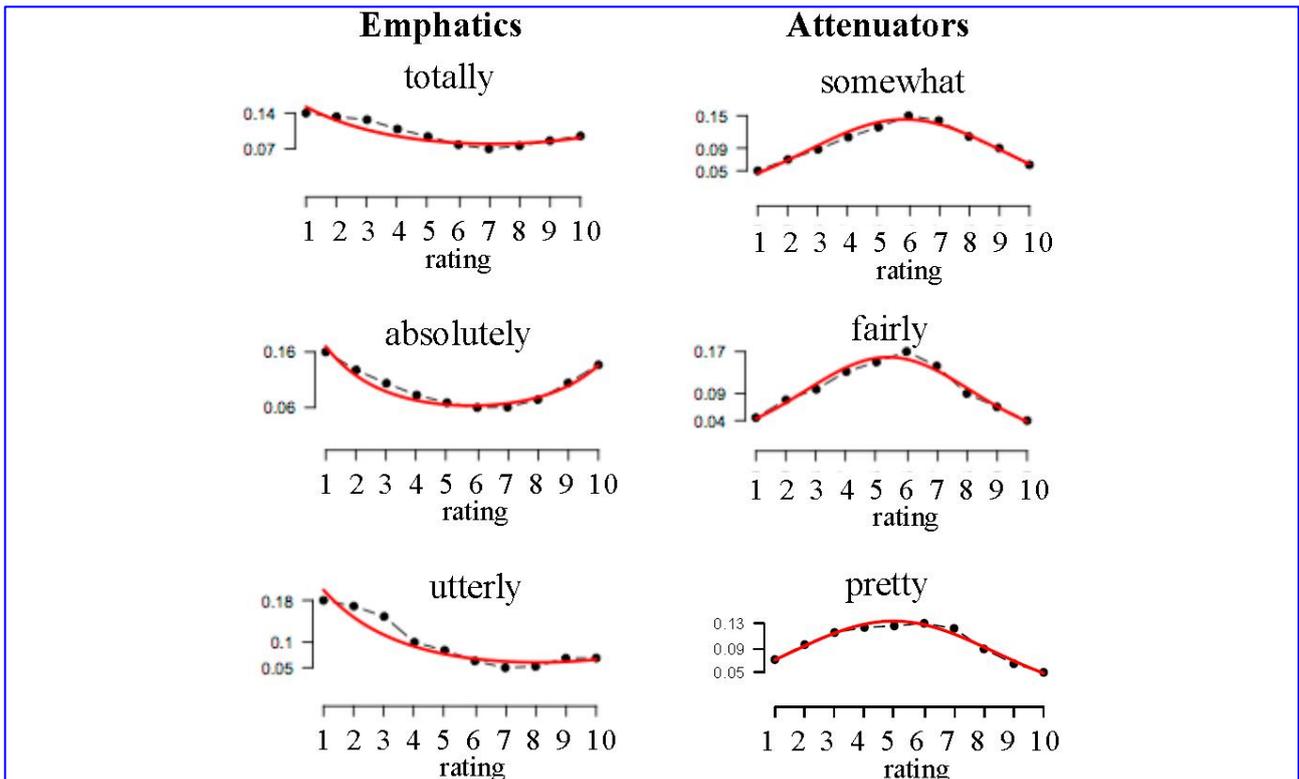


图 20-11: 强调和削弱副词的 Potts 图

图 20.11 显示了用于强调和减弱副词的 Potts 图。请注意，强调倾向于具有 J 形(最有可能出现在最积极的评论中)或 U 形(最有可能出现在强烈的正面和负面情绪中)。衰减器都具有驼峰形状，强调尺度的中部而对两个极端都轻描淡写。这些图既可以用作词汇情感的类型，也可以在情感构成模型中发挥作用。

除了像后验 $P(c|w)$ 、似然 $P(w|c)$ 或归一化似然(公式 20.6)这样的函数外，还使用了许多其他与情感标签一起出现的单词计数的函数。我们将第 20.8 节介绍其中的一些思想，包括像在公式 20.14 中归一化每个作家的计数。

20.5.1. 对数优势比率信息的狄利克雷先验

我们经常想要做的一件事是区分更可能在一类文本中使用的单词，而不是在另一类文本中使用的单词。例如，我们可能想知道与 1 星评价最相关的词和与 5 星评价最相关的词。这些差异可能不仅仅与情感有关。我们可能想要找到民主党议员比共和党议员更经常使用的词语，或者在高档餐厅的菜单上使用的词语比廉价餐厅的多。

给定两类文档，为了找到与某一类比另一类更相关的单词，我们可以测量频率上的差异(单词 w 在 A 类还是 B 类中更频繁?)或者，我们可以计算频率之比，或者计算对数优势比率(两个词之间优势之比的对数)，而不是频率之差。然后，我们可以根据我们选择的任何关联度量对单词进行排序，从类别 A 中过多表示的单词到类别 B 中过多表示的单词。

简单的对数似然或对数优比方法的问题是，它们对非常罕见的单词或非常频繁的单词不起作用;对于频繁使用的单词，所有的差异都显得很大，而对于很少使用的单词，没有什么差异显得很大。

在本节中,我们详解一个解决这个问题的细节:“对数优势比率信息的狄利克雷(Dirichlet)先验”方法(Monroe 等人, 2008), 这是一个特别有用的方法, 用于寻找某些单词, 这些单词与其它单词相比, 在一类特定的文本中统计上的代表性过高。它是基于使用另一个大型语料库来获得我们对每个单词出现频率的预估。

让我们从目标开始: 假设我们想知道 *horrible* 一词在语料库 i 还是在语料库 j 中出现的更多。我们可以计算对数似然比率(log likelihood ratio), 使用 $f^i(w)$ 表示语料库 i 中单词 w 的频率, 而 n^i 表示语料库 i 中单词的总数:

$$\begin{aligned} \text{llr}(\textit{horrible}) &= \log \frac{P^i(\textit{horrible})}{P^j(\textit{horrible})} \\ &= \log P^i(\textit{horrible}) - \log P^j(\textit{horrible}) \\ &= \log \frac{f^i(\textit{horrible})}{n^i} - \log \frac{f^j(\textit{horrible})}{n^j} \end{aligned} \quad (20.7)$$

相反, 让我们计算对数优势比率(log odds ratio): 在 i 或 j 中, *horrible* 的是否有更高的优势:

$$\begin{aligned} \text{lor}(\textit{horrible}) &= \log \left(\frac{P^i(\textit{horrible})}{1 - P^i(\textit{horrible})} \right) - \log \left(\frac{P^j(\textit{horrible})}{1 - P^j(\textit{horrible})} \right) \\ &= \log \left(\frac{\frac{f^i(\textit{horrible})}{n^i}}{1 - \frac{f^i(\textit{horrible})}{n^i}} \right) - \log \left(\frac{\frac{f^j(\textit{horrible})}{n^j}}{1 - \frac{f^j(\textit{horrible})}{n^j}} \right) \\ &= \log \left(\frac{f^i(\textit{horrible})}{n^i - f^i(\textit{horrible})} \right) - \log \left(\frac{f^j(\textit{horrible})}{n^j - f^j(\textit{horrible})} \right) \end{aligned} \quad (20.8)$$

Dirichlet 的直觉是使用一个大型背景语料库来获得我们对每个单词 w 的期望频率的先验估计。我们将通过将语料库中的计数添加到分子和分母中来非常简单地对此操作, 从而实质上是将计数减少到该计数

之前。这就像在询问：给定我们的期望，即在一个良好估计的大背景语料库中给定它们的频率，在此条件下， i 和 j 之间的差异有多大？

该方法通过 w 的先验修正的对数优势比率 $\delta_w^{(i-j)}$ 来估计两个语料库 i 和 j 中单词 w 的频率差异，结果为：

$$\delta_w^{(i-j)} = \log \left(\frac{f_w^i + \alpha_w}{n^i + \alpha_0 - (f_w^i + \alpha_w)} \right) - \log \left(\frac{f_w^j + \alpha_w}{n^j + \alpha_0 - (f_w^j + \alpha_w)} \right) \quad (20.9)$$

(其中 n^i 是语料库 i 的大小， n^j 是语料库 j 的大小， f_w^i 是语料库 i 中单词 w 的计数， f_w^j 是语料库 j 中单词 w 的计数， α_0 是背景语料库的大小，而 α_w 是背景语料库中单词 w 的计数。)

此外，Monroe 等人(2008)利用对数优势比率的方差估计：

$$\sigma^2 \left(\hat{\delta}_w^{(i-j)} \right) \approx \frac{1}{f_w^i + \alpha_w} + \frac{1}{f_w^j + \alpha_w} \quad (20.10)$$

一个单词的最终统计数据是它的对数优势比率的 z 分数：

$$\frac{\hat{\delta}_w^{(i-j)}}{\sqrt{\sigma^2 \left(\hat{\delta}_w^{(i-j)} \right)}} \quad (20.11)$$

Monroe 等人(2008)方法因此以两种方式修改了常用的对数优势比率：使用对数优势比率的 z 分数(控制单词频率中的方差量)，并使用背景语料库中的计数来提供事先单词的计数。

图 20.12 显示了应用于 Yelp 餐厅评论数据集的方法，该方法将 1 星级评论中使用的单词与 5 星级评论中使用的单词进行比较(Jurafsky 等人, 2014)。最大的区别在于明显的情感词，其中 1 星级评论使用诸如 worse, bad, awful 的负面情感词，而 5 星级评论使用诸如 great, best, amazing 的正面情感词。但是，还有其他一些启发性的差异。1 星评价使用逻辑否定(no,not)，而 5 星评价则使用强调并强调普遍性(very, highly, every, always)。1 星评价使用第一人称复数(we,us,our)，而 5 星评价使用第二人称复数。1 星级评论谈论人们(manager,waiter,customer)，而 5 星级评论谈论甜点和昂贵餐厅的属性，例如课程和氛围。参见 Jurafsky 等人(2014)了解更多详情。

Class	Words in 1-star reviews	Class	Words in 5-star reviews
Negative	worst, rude, terrible, horrible, bad, awful, disgusting, bland, tasteless, gross, mediocre, overpriced, worse, poor	Positive	great, best, love(d), delicious, amazing, favorite, perfect, excellent, awesome, friendly, fantastic, fresh, wonderful, incredible, sweet, yum(my)
Negation	no, not	Emphatics/universals	very, highly, perfectly, definitely, absolutely, everything, every, always
1Pl pro	we, us, our	2 pro	you
3 pro	she, he, her, him	Articles	a, the
Past verb	was, were, asked, told, said, did, charged, waited, left, took	Advice	try, recommend
Sequencers	after, then	Conjunct	also, as, well, with, and
Nouns	manager, waitress, waiter, customer, customers, attitude, waste, poisoning, money, bill, minutes	Nouns	atmosphere, dessert, chocolate, wine, course, menu
Irrealis modals	would, should	Auxiliaries	is/'s, can, 've, are
Comp	to, that	Prep, other	in, of, die, city, mouth

图 20-12：与 1 星和 5 星餐厅评论相关的前 50 个单词

图注：使用 Monroe 等人(2008)的方法(Jurafsky 等人, 2014)，在一个有 90 万条评论的 Yelp 数据集中，与 1 星和 5 星餐厅评论相关的前 50 个单词。

20.6. 使用词汇进行情感识别

第四章介绍了用于情感分析的朴素贝叶斯算法。到目前为止，我们在整个章节中所关注的词汇可以用于许多方法来改进情感检测。

在最简单的情况下，当我们没有足够的训练数据来构建监督情感分析器时，可以使用词汇；让人给每个文档分配情感来训练受监督的分类器通常是很昂贵的。

在这种情况下，可以在基于规则的分类算法中使用词汇。最简单的方法就是使用肯定词与否定词的比例：如果一份文档中的肯定词比否定词多(使用词典来确定文档中每个词的极性)，它就被归类为肯定词。通常使用 λ 阈值，其中只有当比值大于 λ 时，文件才被归为正值。如果情感词典中包含每个单词的正权值和负权值 θ_w^+ 和 θ_w^- ，也可以使用。这里有一个简单的情感算法：

$$\begin{aligned}
 f^+ &= \sum_{w \text{ s.t. } w \in \text{positivelexicon}} \theta_w^+ \text{count}(w) \\
 f^- &= \sum_{w \text{ s.t. } w \in \text{negativelexicon}} \theta_w^- \text{count}(w) \\
 \text{sentiment} &= \begin{cases} + & \text{if } \frac{f^+}{f^-} > \lambda \\ - & \text{if } \frac{f^-}{f^+} > \lambda \\ 0 & \text{otherwise.} \end{cases} \quad (20.12)
 \end{aligned}$$

如果有监督训练数据可用，则这些计数从情感词汇计算出来，有时以各种方式加权或归一化，也可以作为特征在分类器中与其他词汇或非词汇特征一起使用。我们将在第 20.8 节回到这些算法。

20.7. 其他任务：人格

许多其他种类的情感含义可以从文本和语音中提取出来。例如，从一个人的语言中检测出他的**人格 (personality)**对于对话系统来说很有用(用户倾向于选择与他们的人格相匹配的施事者)，并且在诸如理解人格如何与其他类型的行为相关联的计算社会科学问题中也很有用。

许多关于人类性格的理论都是基于少数几个维度，例如“五大”维度的不同版本(Digman, 1990)：

外向 vs. 内向： 善于交际、自信、幽默 vs. 冷漠、矜持、害羞

情绪稳定 vs. 神经质： 冷静、不情绪化 vs. 缺乏安全感、焦虑

随和 vs. 不随和： 友好、合作 vs. 敌对、吹毛求疵

认真尽责 vs. 不认真尽责： 自律、有条理 vs. 无效、粗心

经验开放性： 智慧的、富有洞察力的 vs. 肤浅的、缺乏想象力的

通过让作者参加标准的人格测试，一些文本和语音语料库已经被标记为其作者的人格。Pennebaker 和 King(1999)的文章语料库包括 2479 篇来自心理学专业学生的文章(190 万字)，他们被要求在 20 分钟内“写下你想到的任何东西”。梅尔等人(2006)的耳朵(电子激活录音机)文集是通过让志愿者全天佩戴录音机创建的，录音机随机记录下全天的简短对话片段，然后将其转录。Facebook 的语料库(Schwartz 等人, 2013)包括 3.09 亿字的 Facebook 帖子，来自 7.5 万名志愿者。例如，这是 Pennebaker 和 King(1999)的一篇文章的样本，作者属于神经质/情绪稳定量表的神经质一端：

我的一个朋友闯了进来，我从座位上跳了起来。这是疯狂的。我应该告诉他不要再那样做了。其实我并不是那么挑剔。但有些事让我很烦。那些会让我烦恼的事情实际上也会让任何正常人烦恼，所以我知道我不是一个怪物。

而在量表另一端是情感稳定的人：

我应该在这项运动中胜出，因为我知道如何比我认识的任何人更努力地锻炼我的身体，无论什么测试，我总是比其他人更努力地锻炼我的身体。我想成为最好的，无论什么运动或项目。我也应该擅长这个，因为我喜欢骑自行车。

另一种感情含义是 Scherer(2000)所说的人际立场(**interpersonal stance**)，即“在特定的互动中对他人采取的感情立场使人际交往充满色彩”。提取这种含义意味着自动标签参与者是否友好、支持和疏远。例如，Ranganath 等人(2013)研究了一个快速约会的语料库，其中参与者戴着麦克风进行了一系列 4 分钟的浪漫约会。每个参与者都以自己的轻浮、友好、尴尬或自信为自己打标签。Ranganath 等人 (2013)然后结合使用词汇和其他功能从文本中检测出这些人际关系。

20.8. 感情识别

Scherer(2000)所描述的情绪、人格、人际立场和其他感情含义的检测，可以通过概括上述检测情感的算法来完成。

最常见的算法涉及监督分类：为要检测的感情含义标签训练集，并使用从训练集中提取的特征构建分类器。与情感分析一样，如果训练集足够大，并且测试集与训练集足够相似，则只需将所有单词或所有 bigrams 作为功能强大的分类器(如 SVM 或逻辑回归)中的特征，如图 4.2 所示，是一种出色的算法，其性能难以匹敌。因此，我们可以将文本样本的感情含义分类视为简单的文档分类。

然而，对于非常大的数据集，一些修改往往是必要的。例如，Schwartz 等人(2013)利用 Facebook 上 7 亿字的帖子研究了人格、性别和年龄，但只使用了长度为 1-3 的 n-grams 的子集。只有至少 1% 的受试者使用的词和短语被纳入为特征，只有当它们具有足够高的 PMI (PMI 大于 $2 * \text{length}$, length 为单词的个数)时才保留 2-grams 和 3-grams:

$$\text{pmi}(\text{phrase}) = \log \frac{p(\text{phrase})}{\prod_{w \in \text{phrase}} p(w)} \quad (20.13)$$

可以为特征使用各种权重，包括训练集中的原始计数，或一些归一化概率或对数概率。例如，Schwartz 等人(2013)通过对每个主题(subject)的词汇使用总量进行归一化，将特征数转化为短语的似然度。

$$p(\text{phrase} | \text{subject}) = \frac{\text{freq}(\text{phrase}, \text{subject})}{\sum_{\text{phrase}' \in \text{vocab}(\text{subject})} \text{freq}(\text{phrase}', \text{subject})} \quad (20.14)$$

如果训练数据稀疏或与测试集不同，那么我们讨论的任何词汇都可以发挥有益的作用，无论是单独使用还是与所有单词和 n-gram 结合使用。许多可能的值可用于词汇功能。最简单的只是一个指示符函数，如果特定文本中有来自相关词汇 L 的任何单词，则特征 f_L 的值将取值为 1。使用第 4 章的符号，其中为特定的输出类 c 和文档 x 定义了一个特征值。

$$f_L(c, x) = \begin{cases} 1 & \text{if } \exists w : w \in L \ \& \ w \in x \ \& \ \text{class} = c \\ 0 & \text{otherwise} \end{cases}$$

或者，某个特定词汇 L 的特征 f_L 的值可以是出现在 L 中的文档中单词符记的总数:

$$f_L = \sum_{w \in L} \text{count}(w)$$

对于每个单词都与得分或权重相关联的词汇，可以将计数乘以权重 θ_w^L :

$$f_L = \sum_{w \in L} \theta_w^L \text{count}(w)$$

也可以像等式 20.14 中那样按作者对计数取对数或归一化。

无论它们如何定义，这些词汇特征都将在监督分类器中用于预测文本或文档的所需感情类别。训练完分类器后，我们可以检查哪些词汇功能与哪些类相关联。对于逻辑回归等分类器，特征权重可指示特征与

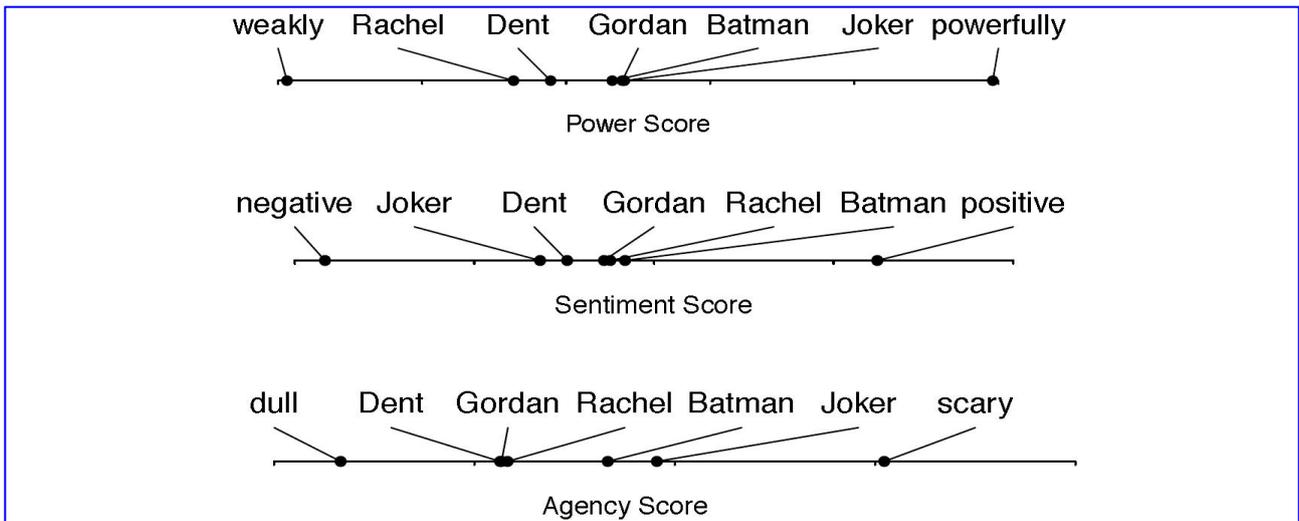


图 20-14: 《黑暗骑士》中角色的权力(支配)、情感(效价)和施事者(唤起)

图注: 是通过 NRC VAD 词典训练的嵌入计算出来的。注意主角(蝙蝠侠)和对手(小丑)的权力和施事者得分很高, 但情感上有所不同, 而爱慕对象 Rachel 的权力和施事者都很低, 但情感很高。

考虑这样的句子:

(20.15) Country A violated the sovereignty of Country B

(20.16) the teenager ... survived the Boston Marathon bombing

通过使用(20.15)中的动词 *violate*, 作者表达了对国家 B 的同情, 将国家 B 描绘为受害者, 并对施事者国家 A 表示敌对。相反, 使用 *survive* 这个动词时, 作者(20.16)表示轰炸是一种消极的经历, 而句子的主题(少年)是一个同情的角色。这些内涵的体是动词 *violate* 和 *survive* 的意义所固有的, 见图 20.15。

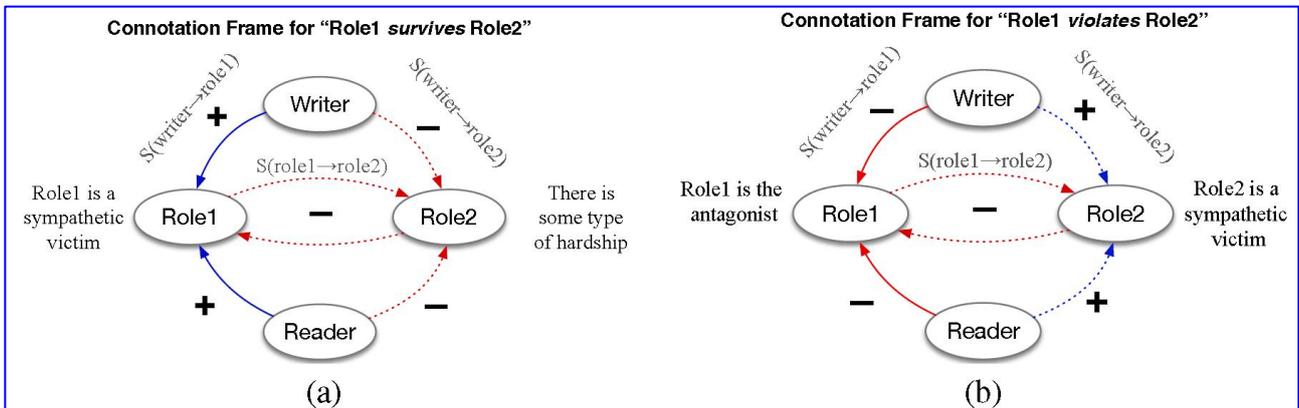


图 20-15: *survive* 和 *violate* 的内涵框架

图注: (a) *survive*: 作者和读者对角色 1(主语)有积极的情感, 对角色 2(直接宾语)有消极的情感。

(b) *violate*: 作者和读者对角色 2(直接宾语)有积极的情感。

Rashkin 等人(2016)和 Rashkin 等人(2017)的内涵框架词汇也表达了谓词对每个论元的其他内涵方面, 包括两个方面, 第一是效果(x 发生了不好的事情)值: (x 是有价值的); 第二是心理状态: (x 被事件困扰)。内涵框架也可以标记论元之间的权力差异(使用动词 *implore* 意味着主题论元比施事者更有权力), 和每个论元的施事者(*waited* 是低级施事者)。图 20.16 显示了 Sap 等人(2017)的可视化结果。

可以手动构建内涵框架(Sap 等人, 2017), 也可以通过监督学习(Rashkin 等人, 2016)来学习它们, 例如使用手工标签的训练数据来监督每个个体关系的分类器, 例如 $S(\text{writer} \rightarrow \text{Role1})$ 是+还是-, 然后通过所有关系上的全局约束来提高准确性。

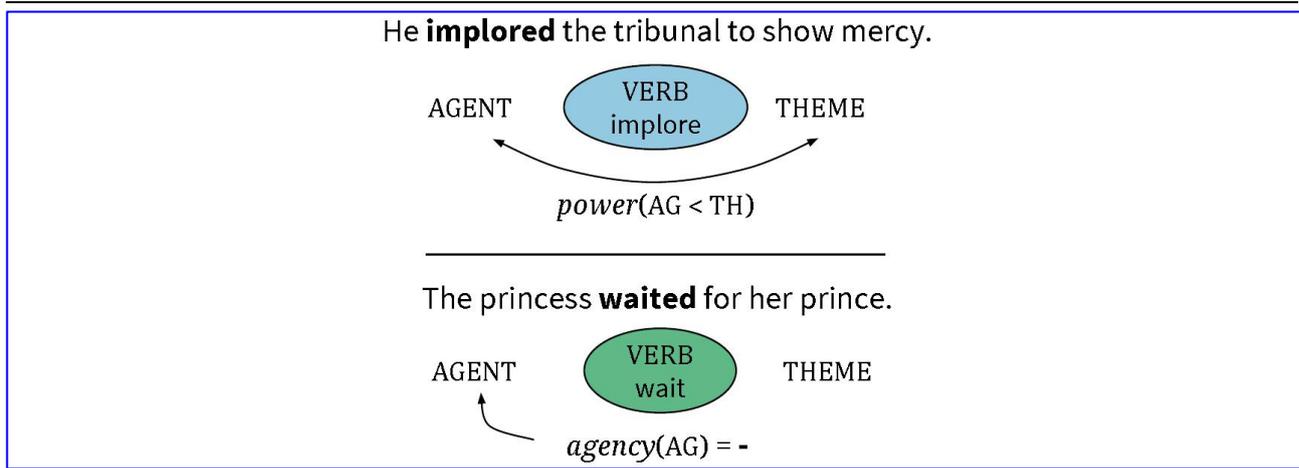


图 20-16: Sap 的内涵框架

图注：图中动词 implore 意味着施事者的权力低于主题(对照于动词 demanded)，表明 waited 的主语的施事者水平较低。

20.11. 总结

- 可以区分多种**感情状态**，包括：情绪、语气、**态度(包括情感)**、人际立场和性格。
- **情绪**可用固定的原子单位表示，常被称为基本情绪，也可表示为**效价**和**唤起**等维度定义的空间中的点。
- 单词具有与这些感情状态有关的**内涵**体，单词含义的这一内涵体可以用词汇表示。
- 可以人工创建感情词汇，使用**众包**来标签每个单词的感情内容。
- 可以使用相似度度量(嵌入余弦)从种子词的自举引导过程中建立**半监督**的词汇。
- 在世界范围内可以找到方便的训练信号(例如，用户在查看站点上分配的评分)时，可以以**全监督**的方式学习词汇。
- 通过在训练文本中使用单词计数的各种函数，以及像**对数优势比率信息 Dirichlet 先验**这样的比率指标，单词可以在词汇中分配权重。
- **性格**通常表现为 5 维空间中的一个点。
- 感情与情感一样，可以通过使用标准的监督**文本分类**技术，将文本中的所有单词或字母作为特征来检测。附加特征可以从词汇表中的单词计数中抽取。
- 词汇也可用于检测**基于规则**的分类器中的感情，根据每个词汇表中单词的数量选择简单多数情感。
- **内涵框架**表达了谓词对其论元所编码的更丰富的感情含义关系。

20.12. 文献和历史说明

正式表达词语主观意义的想法始于 Osgood 等人(1957)，这一开创性研究首次提出了第 6 章中描述的含义向量空间模型。Osgood 等人(1957)让参与者在不同的量表上对单词进行评分，并对评分进行因子分析。他们发现的最重要的因素是评价维度，它区分了好/坏、有价值/无价值、令人愉快/不愉快的配对。这部作品影响了内容分析领域早期情感和感情含义字典的发展(Stone 等人，1966)。

Wiebe(1994)在检测文本主观性(subjectivity)方面开始了一项有影响的工作，首先是识别主观句子和文本中被描述为持有私人状态、信仰或态度的主观人物。Hatzivassiloglou 和 McKeown(1997)的极性词汇等习得情感词汇在主观性检测中被证明是一个有用的特征(Hatzivassiloglou 和 Wiebe 2000, Wiebe 2000)。

情感一词似乎是 Das 和 Chen(2001)在 2001 年引入的，用来描述通过查看股票交易留言板上的词汇来衡量市场情感的任务。Das 和 Chen(2001)在同一篇论文中也提出了情感词汇的使用。词汇中的单词列表是手工创建的，但是每个单词都通过最大化类间差异和最小化类内差异来分配权重，这取决于它对特定类的区别程度(比如 buy 和 sell)。情感一词以及词汇的使用很快就流行起来了(例如，Turney 在 2002 年的一篇文章中就提到了这一点)。Pang 等人(2002)首先展示了在没有情感词汇的情况下使用所有单词的能力；参见 Wang 和 Manning(2012)。

我们用来扩展情感字典的大部分半监督方法都是基于同义词和反义词在同一句子中往往同时出现的早期想法(Miller 和 Charles 1991, Justeson 和 Katz 1991, Riloff 和 Shepherd 1997)。其他学习感情含义线索的半监督方法依赖于信息提取技术，如 AutoSlog 模式提取器(Riloff 和 Wiebe, 2003)。Hatzivassiloglou 和 McKeown(1997)首先提出了基于图的情感算法，而图传播成为了标准方法(Zhu 和 Ghahramani 2002, Zhu 等人 2003, Zhou 等人 2004, Velikovich 等人 2010)。众包也可以通过过滤半监督词汇学习的结果来提高精度(Riloff 和 Shepherd 1997, Fast 等人 2016)。

最近的许多工作集中于学习直接编码情感或其他属性的嵌入的方法，例如 Rothe 等人(2016)的 DENSIFIER 算法，该算法学习如何转换嵌入空间，专注于情感(或其他)信息。

20.13. 练习

(无)