



自然语言处理导论

张奇 桂韬 黄萱菁

2022 年 12 月 12 日

数与数组

α	标量
$\boldsymbol{\alpha}$	向量
\boldsymbol{A}	矩阵
\mathbf{A}	张量
\boldsymbol{I}_n	n 行 n 列单位矩阵
\boldsymbol{v}_w	单词 w 的分布式向量表示
\boldsymbol{e}_w	单词 w 的独热向量表示: $[0,0,...,1,0,...0]$, w 下标处元素为 1

索引

α_i	向量 $\boldsymbol{\alpha}$ 中索引 i 处的元素
$\boldsymbol{\alpha}_{-i}$	向量 $\boldsymbol{\alpha}$ 中除索引 i 之外的元素
$w_{i:j}$	序列 w 中从第 i 个元素到第 j 个元素组成的片段或子序列
A_{ij}	矩阵 \boldsymbol{A} 中第 i 行、第 j 列处的元素
\boldsymbol{A}_i	矩阵 \boldsymbol{A} 中第 i 行
$\boldsymbol{A}_{:j}$	矩阵 \boldsymbol{A} 中第 j 列
A_{ijk}	三维张量 \mathbf{A} 中索引为 (i, j, k) 处元素
$\mathbf{A}::i$	三维张量 \mathbf{A} 中的一个二维切片

集合

\mathbb{A}	集合
\mathbb{R}	实数集合
$\{0, 1\}$	含 0 和 1 的二值集合
$\{0, 1, ..., n\}$	含 0 和 n 的正整数的集合
$[a, b]$	a 到 b 的实数闭区间
$(a, b]$	a 到 b 的实数左开右闭区间

线性代数

\mathbf{A}^\top	矩阵 \mathbf{A} 的转置
$\mathbf{A} \odot \mathbf{B}$	矩阵 \mathbf{A} 与矩阵 \mathbf{B} 的 Hardamard 乘积
$\det \mathbf{A}^\top$	矩阵 \mathbf{A} 的行列式
$[\mathbf{x}; \mathbf{y}]$	向量 \mathbf{x} 与 \mathbf{y} 的拼接
$[\mathbf{U}; \mathbf{V}]$	矩阵 \mathbf{A} 与 \mathbf{V} 沿行向量拼接
$\mathbf{x} \cdot \mathbf{y}$ 或 $\mathbf{x}^\top \mathbf{y}$	向量 \mathbf{x} 与 \mathbf{y} 的点积

微积分

$\frac{dy}{dx}$	y 对 x 的导数
$\frac{\partial y}{\partial x}$	y 对 x 的偏导数
$\nabla_{\mathbf{x}} y$	y 对向量 \mathbf{x} 的梯度
$\nabla_{\mathbf{X}} y$	y 对矩阵 \mathbf{X} 的梯度
$\nabla_{\mathbf{x}} y$	y 对张量 \mathbf{X} 的梯度

概率与信息论

$a \perp b$	随机变量 a 与 b 独立
$a \perp b \mid c$	随机变量 a 与 b 关于 c 条件独立
$P(a)$	离散变量概率分布
$p(a)$	连续变量概率分布
$a \sim P$	随机变量 a 服从分布 P
$\mathbb{E}_{x \sim P}[f(x)]$ 或 $\mathbb{E}[f(x)]$	$f(x)$ 在分布 $P(x)$ 下的期望
$\text{Var}(f(x))$	$f(x)$ 在分布 $P(x)$ 下的方差
$\text{Cov}(f(x), g(x))$	$f(x)$ 与 $g(x)$ 在分布 $P(x)$ 下的协方差
$H(f(x))$	随机变量 x 的信息熵
$D_{KL}(P \parallel Q)$	概率分布 P 与 Q 的 KL 散度
$\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	均值为 $\boldsymbol{\mu}$ 、协方差为 $\boldsymbol{\Sigma}$ 的高斯分布

数据与概率分布

\mathbb{X}	数据集
$\mathbf{x}^{(i)}$	数据集中第 i 个样本（输入）
$\mathbf{y}^{(i)}$ 或 $y^{(i)}$	第 i 个样本 $\mathbf{x}^{(i)}$ 的标签（输出）

函数

$f : \mathcal{A} \longrightarrow \mathcal{B}$	由定义域 \mathcal{A} 到值域 \mathcal{B} 的函数（映射） f
$f \circ g$	f 与 g 的复合函数
$f(\boldsymbol{x}; \boldsymbol{\theta})$	由参数 $\boldsymbol{\theta}$ 定义的关于 \boldsymbol{x} 的函数（也可以直接写作 $f(\boldsymbol{x})$ ，省略 $\boldsymbol{\theta}$ ）
$\log x$	x 的自然对数函数
$\sigma(x)$	Sigmoid 函数 $\frac{1}{1 + \exp(-x)}$
$\ \boldsymbol{x}\ _p$	\boldsymbol{x} 的 L^p 范数
$\ \boldsymbol{x}\ $	\boldsymbol{x} 的 L^2 范数
$\mathbf{1}^{\text{condition}}$	条件指示函数：如果 condition 为真，则值为 1；否则值为 0

本书中常用写法

- 给定词表 \mathbb{V} ，其大小为 $|\mathbb{V}|$
- 序列 $x = x_1, x_2, \dots, x_n$ 中第 i 个单词 x_i 的词向量 \boldsymbol{v}_{x_i}
- 损失函数 \mathcal{L} 为负对数似然函数： $\mathcal{L}(\boldsymbol{\theta}) = -\sum_{(x,y)} \log P(y|x_1 \dots x_n)$
- 算法的空间复杂度为 $\mathcal{O}(mn)$

目 录

9 情感分析	1
9.1 情感分析概述	1
9.1.1 情感模型	2
9.1.2 情感分析主要任务	5
9.2 篇章级情感分析	9
9.2.1 基于支持向量机的篇章级情感分析	9
9.2.2 基于层次结构的篇章级情感分析	12
9.2.3 篇章级情感分析语料库	14
9.3 句子级情感分析	16
9.3.1 基于词典的句子级情感分析	16
9.3.2 基于递归神经张量网络的句子级情感分析	17
9.3.3 基于情感知识增强预训练的句子级情感分析	19
9.3.4 句子级情感分析语料库	21
9.4 属性级情感分析	22
9.4.1 情感信息抽取	22
9.4.2 属性级情感分类	28
9.4.3 属性级情感分析语料库	38
9.5 延伸阅读	40
9.6 习题	41

9. 情感分析

人类在语言交流过程中通常富含丰富的情感信息, 如何自动理解人类语言信息中的情感, 是自然语言处理领域的研究热点与难点。情感分析 (Sentiment analysis) 又称观点挖掘 (Opinion Mining), 其任务旨在从文本中分析得到人们关于主题或实体的评价、观点或态度, 也包括分析文本所表达的情绪信息。情感分析包含了评论挖掘、评价抽取、主观性分析、情绪分析等多种不同的任务, 在舆情分析、情报挖掘、电子商务、对话系统等领域具有广泛的应用。近年来, 随着社交媒体快速发展, 用户通过各种平台发表和分享大量评论和观点内容, 情感分析技术无论是在研究还是应用上均取得了显著的进展。

本章首先介绍情感分析的基本概念, 在此基础上根据语言单元的颗粒度划分, 分别介绍篇章级、句子级以及属性级情感倾向分析任务和常用算法。

9.1 情感分析概述

随着互联网的快速发展, 特别是以博客、微博为代表的 Web 2.0 平台普及, 用户在各类社会媒体平台上产生了大量包含评论和观点的文本内容。通过对这些内容进行分析, 可以有效的了解用户喜好、发现产品需求、市场情绪分析。大量应用需求也促使自 2000 年以来, 情感分析任务逐渐受到越来越多学术界和产业界的关注, 包括倾向性形容词发现^[1]、股票市场情绪分析^[2]、观点分析与跟踪^[3]、评论倾向性分析^[4]等任务相继开展。2003 年, 情感分析和观点挖掘的这两个术语也相继由 Nasukawa 等人 [5] 和 Dave 等人 [6] 提出。更早的一些工作包括篇章中主观性分析^[7]、观点跟踪^[8]、主观性分类^[9]等任务在 20 世纪 90 年代起也已经开始。

情感分析包含的研究内容众多, 包括自然语言处理、数据挖掘、机器学习等不同领域研究人员都对该任务开展了大量研究, 也因此造成情感分析相关术语繁杂, 有很多研究内容大体相同但是又有一些微小区别的任务, 刘兵教授在其关于情感分析专著中^[10]给出的任务包括: 情感分析 (sentiment analysis)、观点挖掘 (opinion mining)、观点抽取 (opinion extraction)、情感挖掘 (sentiment mining)、主客观分析 (subjectivity analysis)、感情分析 (affect analysis)、情绪分析 (emotion analysis)、评论挖掘 (review mining) 等。在本书中, 我们采用刘兵教授的建议, 使用情感分析来代表该领域整体研究内容。

本节首先对情感模型进行简单介绍，在此基础上对情感分析主要研究内容进行介绍，包括情感分类情感单元抽、情绪分类、观点摘要析等。

9.1.1 情感模型

根据情感分析任务定义，我们可以看到情感分析主要包含两个主要任务：1) 分析文本中针对某个主题或实体的评价或观点；2) 分析文本中所表达的情绪类的情感。虽然针对观点分析和情绪分析任务所采用的自然语言处理算法非常类似，但是在语言学 and 心理学理论中，上述问题还是有非常大的不同。本节将介绍常见的观点类和情绪类理论模型。

1. 观点模型

观点 (opinion) 是指从某一立场或角度出发对事物所持的看法或态度，是一种表达了感觉、看法、信念的陈述。观点的情感倾向也称为极性，可以是正面 (褒义)、负面 (贬义) 或中立。每种不同的情感倾向还具有不同的强度。比如“完美”比“好”表达的褒义程度更强。很多副词也具有增强或者减弱情感倾向的作用。常见的增强词包括：“很”、“非常”、“very”、“extremely”等。常见的减弱词包括：“可能”、“一定程度上”、“slightly”、“a little bit”等。可以采用情感评分的方法，使用离散化的评分表达情感的强度。比如，可以将情感分为 5 档 (1-5 分)，1 分表示强烈负面，2 分表示负面，3 分表示中立，4 分表示正面，5 分表示强烈正面。观点从语言学、心理学等不同角度划分为不同类型，包括：常规型观点和比较型观点、显式观点和隐式观点、感性观点和理性观点等。

常规观点通常简称为观点，是指通过直接或间接的方式对事物所表述的观点^[11]。例如：

(1) 这家餐厅非常差。

(2) 更换了这台显示器后，我的眼睛感觉非常舒服。

上例中，两个句子都属于常规性观点，句子 (1) 是直接观点，直接针对餐厅表述了句子作者的负面评价；句子 (2) 是间接观点，通过描述眼睛的感受，间接的表达了对显示器的正面评价。通常情况下，直接观点出现的比例相较间接观点高很多。目前的工作大多针对直接观点开展，间接观点的研究相对较少。

比较型观点 (Comparative Sentiment) 是指对两个或更多事物之间的相同或者不同点进行比较，并表达了观点持有者对其中一个事物的态度^[12]。例如：

(1) 这家餐厅的环境比人民路上那家的好很多。

(2) 新一代的显卡的显存相较于上一代有了大幅度的提升。

上例中，两个句子都属于典型比较型观点。两句中，虽然都没有对实体直接给出评价，但是对实体在某个属性上给出排序。在英语中，比较型观点通常或通过形容词或者副词的比较级或者最高级进行表达，但也有一些例外 (如 prefer)。

人们对观点的表达通常使用包含情感词语的主观性句子，但是也有一些观点的表达通过客观事实描述性句子完成。基于此，可以划分为显式观点和隐式观点两类。显式观点 (Explicit Sentiment) 是指在句子中通过常规观点或比较型观点句子直接对观点、看法、感受进行表达的句子。例如：

- (1) 不错, 交通便利, 方便出行!
- (2) 展馆太小了, 场景少, 一般般。

上例中, 观点持有者直接表达了自己的感受和评价, 都属于典型的显式观点。

隐式观点 (Implicit Sentiment) 是指在客观或者真实的事实陈述中蕴含的常规型或比较型的观点。例如:

- (1) 该款电动车续航里程可达 1000 公里。
- (2) 雪山脚下的一个景点, 从进门到出去给了半个小时游览。

上例中, 两个句子都是在描述事实, 虽然没有使用包含情感的词语, 但是也表达了观点持有者的态度。隐式观点虽然在观点句中的占比不高, 但是由于是否包含观点以及观点极性分类通常需要依赖常识, 使得隐式观点的识别和分类的难度很高。上例中句子 (1) 表达了正面的观点, 这需要依赖当前电动车续航里程通常小于 600 公里这一背景知识。

理性情感 (Rational Sentiment) 是指观点来源于理性的推理, 不包含主观情绪。例如:

- (1) 笔记本电池不错, 可以连续使用 40 小时。
- (2) 酒店环境很好, 距离沙滩仅有 200 米。

上例中, 虽然都包含主观性表达的观点, 但都是根据相对理性的方式以客观事实为依据的理性推理。

感性情感 (Emotional Sentiment) 是指观点来源于观点持有者的感性的主观表达。例如:

- (1) 这个酒店太垃圾了。
- (2) 这是最好的车。

与理性情感相比, 感性情感通常更加强烈, 在整体评论中的数量也更多。

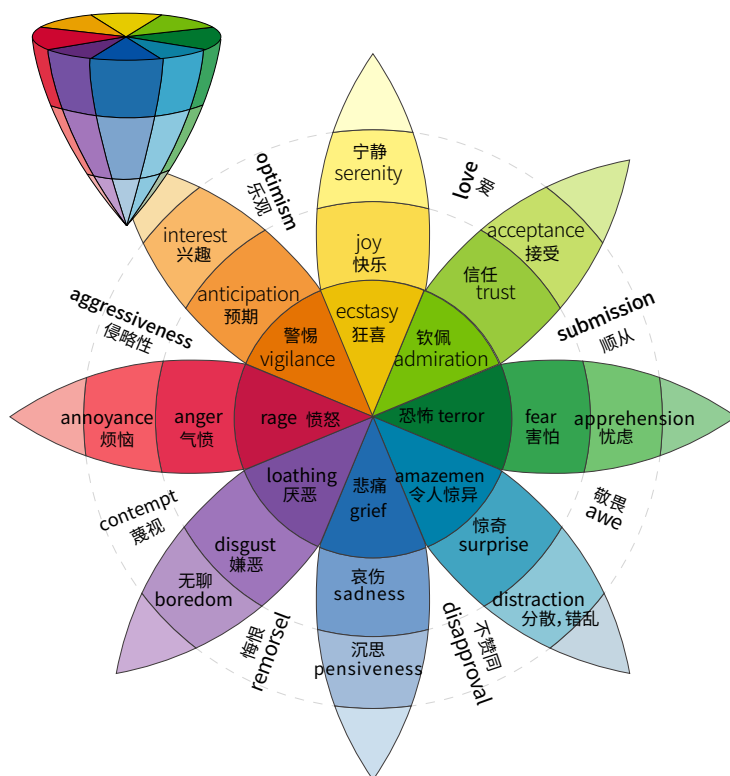
2. 情绪模型

情绪 (emotion) 是人们对外界刺激所产生的反映, 由多种感觉、思想以及行为综合产生的生理和心理状态。包括喜爱、欢乐、悲伤等。由于自然语言的复杂性和人类情绪的多边形, 不同领域的研究学者对情绪类别的划分也有不同。我国古代《礼记》中就对人的情绪进行了“七情”划分: 喜、怒、哀、惧、爱、恶、欲。心理学的理论学者 Parrott 把情绪进行了更细粒度的花费, 不仅仅给出了基本情绪, 还给出了二级乃至三级等更细粒度的情绪 [13] (如表9.1所示)。

2001 年, 心理学家 Plutchik 基于进化规则的综合理论, 提出了多维度情绪模型^[14]。该模型定义了 8 种基本基本情绪, 分为 4 对双向组合: 高兴与悲伤、愤怒与恐惧、信任与厌恶、诧异与期望 (joy vs. sadness, anger vs. fear, trust vs. disgust, surprise vs. anticipation)。图9.1给出了 Plutchik 模型的情绪类别在“情绪轮”上的排序, 其中颜色深浅代表这种情绪的饱和度, 离圆心的远近代表情绪的强度。在 Plutchik 情绪理论中, 每种情绪都可以分为 3 度, 例如, 宁静 (serenity) 是最小程度的高兴, 是不饱和状态; 狂喜 (ecstasy) 是最高程度的高兴, 是饱和状态。此外, Plutchik 还提出一种假设, 认为两种相邻的基本情绪可以组合成一种符合情绪, 例如, 快乐 (joy) + 预期 (anticipation) = 乐观 (optimism)。

基本情绪	二级情绪类型	三级情绪类型
Anger	Disgust	Contempt, loathing, revulsion
	Envy	Jealousy
	Exasperation	Frustration
	Irritability	Aggravation, agitation, annoyance, crosspatch, grouchy, grumpy
	Rage	Anger, bitter, dislike, ferocity, fury, haterd, hostility, outrage, resentment, scorn, spite, vengefulness, wrath
	Torment	Torment
Fear	Horror	Alarm, fear, fright, horror, hysteria, mortification, panic, shock, terror
	Nervousness	Anxiety, apprehension(fear), distress, dread, suspense, uneasiness, worry
Joy	Cheerfulness	Amusement, bliss, gaiety, glee, jolliness, joviality, joy, delight, enjoyment, gladness, happiness, jubilation, elation, satisfaction, ecstasy, euphoria
	Contentment	Pleasure
	Enthrallment	Enthrallment, rapture
	Optimism	Eagerness, hope
	Pride	Triumph
	Relief	Relief
	Zest	Enthusiasm, excitement, exhilaration, thrill, zeal
Love	Affection	Adoration, attractiveness, caring, compassion, fondness, liking, sentimentality, tenderness
	Longing	Longing
	Lust/sexual desire	Desire, infatuation, passion
Sadness	Disappointment	Dismay, displeasure
	Neglect	Alienation, defeatism, dejection, embarrassment, homesickness, humiliation, insecurity, insult, isolation, loneliness, rejection
	Sadness	Depression, despair, gloom, glumness, grief, melancholy, misery, sorrow, unhappy, woe
	Shame	Guilt, regret, remorse
	Suffering	Agony, anguish, hurt
	Sympathy	Pity, sympathy
Surprise	Surprise	Amazement, astonishment

表 9.1 基本情绪类型、二级情绪和三级情绪^[13].

图 9.1 Plutchik 提出的情绪轮^[14]

除了上述基于类别空间的情绪模型外，还有一些情绪模型是基于维度空间思想。这类情绪模型认为情绪并不是相互分离和独立的，而是相互联系和交叉的，某种特定的情绪可以表示为连续维度空间的中的点或向量。Mehrabian 与 Russell 提出了 PAD 三维模型 (pleasure-arousal-dominance) [15]，提出情绪可以从愉悦度 (pleasure)、唤醒度 (arousal) 和支配度 (dominance) 进行分解。此后，Russell 进一步对 PAD 理论进行了修证^[16]，提出愉悦度和唤醒两个维度就能够解释绝大部分的情绪，支配度更多与认知活动有关，因此就有了由愉悦度和唤醒度构成情绪环结构。图9.2给出了根据 PAD 理论和 PA 理论表示的情绪。

9.1.2 情感分析主要任务

情感分析包含非常类型各种任务，从任务的类型层面可以划分为情感分类和情感信息抽取两大类。从语言单元层面又可以划分为篇章级情感分析 (Document-Level Sentiment Analysis)，句子级情感分析 (Sentence-Level Sentiment Analysis) 和属性级情感分析 (又称方面级情感分析, Aspect-Level Sentiment Analysis) ^[17, 18]。本节中将按照情感分类和情感信息抽取分别介绍情感分析主要任

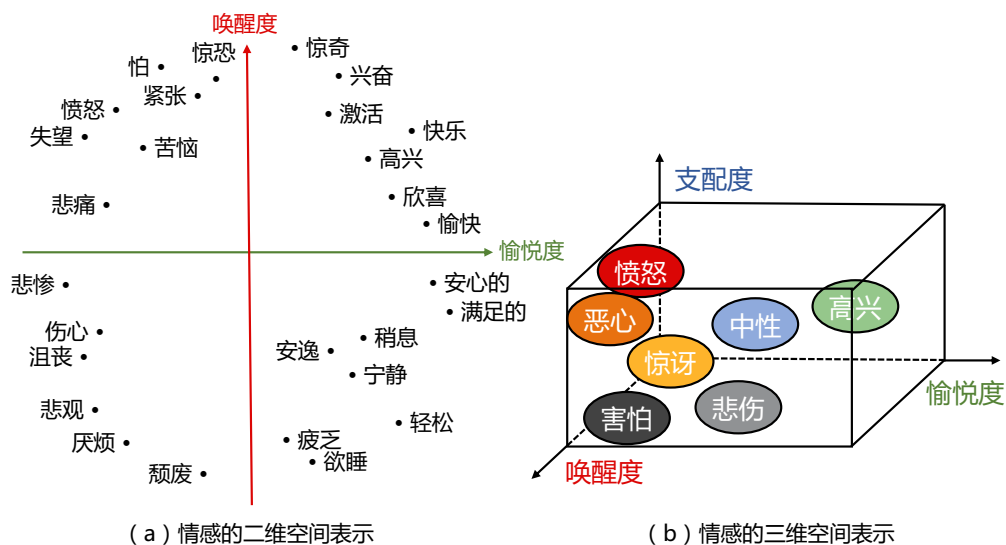


图 9.2 基于维度空间表示方法

务，并对情感摘要、虚假观点检测等情感相关任务进行简要介绍。

1. 情感分类任务

情感分类任务（Sentiment Classification）的目标是根据给定文本内容，识别所蕴含的情感或观点，并确定情感的类别或观点倾向性。主要任务包括主客观识别、情感极性判断、情感强度判断、情绪分类等。

主客观分类（subjective classification），又称观点识别（opinion identification），目标是判断给定的文本是否包含情感或观点，即判断文本是主观性（subjective）还是客观性（objective）。例如：

- (1) 味道不错！团购很实惠。
- (2) 复旦大学校名取自《尚书大传》之“日月光华，旦复旦兮”。

上例中，句子 (1) 是主观性句子，表达了用户的观点；句子 (2) 是客观性句子，没有表达任何观点或态度。

极性分类（polarity classification）目标是判断给定的文本情感或观点的情感极性，即判断文本的情感是正面（褒义）、负面（贬义）还是中性。例如：

- (1) 环境相当不错，业务水平很专业。
- (2) 实在是很坑的一个景区。
- (3) 地理位置也还可以。

上例中，句子 (1) 是正面评价；句子 (2) 是负面评价；句子 (3) 是中性评价。需要特别说明的是中性评价不等同于客观性文本，只是句子所表达的情感极性并不能归于正面和负面。

情绪分类 (emotion classification) 目标是判断给定的文本蕴含的情绪类别, 即判断文本中情绪是喜、怒、哀、恶等。情绪类别需要根据在上节中介绍的情绪模型进行选择。例如:

(1) 我的心里绽开了朵朵鲜花, 就要蹦出来似的。

(2) 钟表, 可以回到起点, 却已不是昨天。

上例中, 句子 (1) 是表达了快乐的情绪; 句子 (2) 表达了悲伤的情绪。

情感强度判别 (sentiment strength detection) 目标是判断给定文本的情感强度, 即判断文本的情感是强烈正向、正向、中性、负向、强烈负向等。根据上节中介绍的观点模型, 也可以采用分数表示情感强度。例如:

(1) 这地方交通不太方便了。

(2) 这地方交通实在是太不方便了。

上例中, 句子 (1) 表达了负面的情感; 句子 (2) 表达了强烈的负面情感。情感强度判断可以采用分类方法进行建模, 也可以转化为回归问题进行解决。

根据语言单元粒度不同, 上述情感分类任务可以划分为篇章级、句子级和属性级, 即篇章级主客观分类、句子级情绪分类、属性级情感强度判别等。篇章级和句子级的主要不同在于所处理的文本粒度不同。而属性级的情感分类任务所处理的文本可能是篇章或者句子, 但是目标并不是判断整个句子或者篇章整体的情感, 而是判断文本内容中关于该属性的情感类别或情感极性。因此, 属性级的情感分类任务, 不仅要输入文本内容, 还要输入所关注的属性单元。

2. 情感信息抽取任务

情感信息抽取 (Sentiment Information Extraction), 也称**评价要素抽取 (Opinion Elements Extraction)**, 目标是抽取文本的中表达情感的核心要素, 如评价词、评价对象、观点持有者、评价搭配等。相较于情感分类任务, 情感信息抽取可以获得结构化的情感信息。情感信息抽取任务也与属性级情感分类任务有紧密联系, 可以在情感抽取的基础上完成属性级观点识别等任务, 利用抽取出的评价词以及评价搭配使得属性级观点识别具有一定的可解释性。

例如: [懂车会]观点持有者: [车窗]评价对象 采用无边框玻璃设计, [很酷]评价词, 但吹毛求疵一点的, 会觉得 [隔音]评价对象 [不算太好]评价词。

评价对象抽取 (opinion target extraction) 目标是抽取文本中的被评价对象的主体。上例中, “车窗”和“隔音”都属于评价对象。

评价词抽取 (opinion word extraction) 目标是抽取文本中所使用的评价词。上例中, “很酷”和“不算太好”都属于评价词。

评价搭配抽取 (opinion collocation extraction) 目标是识别文本中评价对象所对应的评价词。可以使用二元组表示 < 评价对象, 评价词 >。上例中, < 车窗, 很酷 > 以及 < 隔音, 不算太好 > 都属于评价搭配。

评价搭配极性判别 (opinion collocation polarity classification) 目标是判断某个评价搭配的情感极性。上例中, “车窗”与“很酷”是正面评价, “隔音”与“不算太好”是负面评价。

观点持有者（opinion holder extraction）目标是抽取文本中观点的持有者。上例中，“懂车会”作为作者给出了上述评价，属于观点持有者。大部分情况下观点持有者是文章的作者，不一定体现在当前的文本中。还有一些文章中引用了他人的评价，观点持有者就应该是所引用的内容的作者。

3. 情感相关其他任务

除了情感分类任务和情感信息抽取任务之外，还有一些任务与情感和观点相关，主要包括：观点摘要、辩论立场检测、评论质量判断、虚假观点检测等。

观点摘要（Opinion Summarization）是以评价对象以及以针对评价对象的的观点为中心进行的单文档或多文档摘要。观点摘要不仅可以与普通文本摘要类似，由抽取的重要句子或生成的文本构成摘要，还可以采用情感信息抽取获得的结构化数据进行综合后可可视化的形式。图9.3给出了基于属性的观点摘要样例 [19]，展示了两个数码相机在各属性上的评价对比。

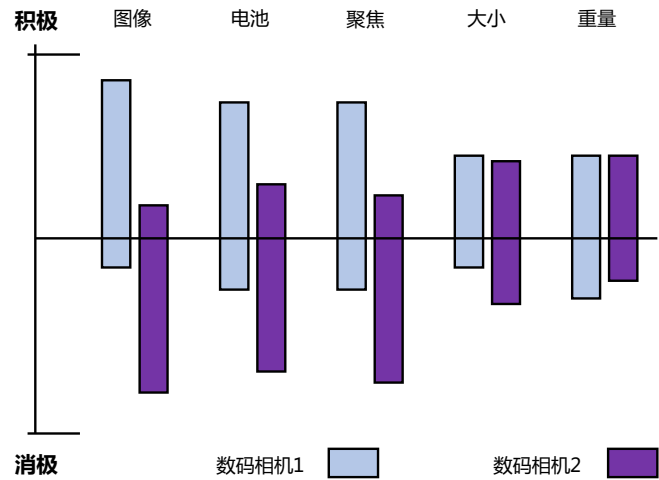


图 9.3 基于属性的观点摘要样例

社交媒体中不仅包含用户发表的评论信息，还包含大量的参与者之间存在交互的辩论、讨论或评论。辩论立场检测（Stance Detection）目标是从辩论文本中识别用户对于某个辩论主题的立场，即用户是支持这个主题还是反对这个主题。这项工作与属性级关系分析较为类似，但是不同是产品的属性通常是预先定义的有限集合，而论辩主题多种多样，通常无法预先定义。

例如：爆竹声中一岁除，春风送暖入屠苏。”春节燃放烟花爆竹，自古以来就是中国人的传统习俗。不过近几年来，随着“双禁”的规定越发严格，过年没有鞭炮声，过年越来越没有年味儿了。

上例中，用户表达了对“春节不应该放鞭炮”议题反对的立场。

随着用户对社交媒体中的评论的重视，一些组织和个人在平台中发布虚假评论信息，以达到

抹黑某个产品或者宣传某个产品的目的。这些虚假评论也称为垃圾评论（Spam Review）。如果不能甄别这些垃圾评论，会对社交媒体充满虚假和谎言，对于平台公信力造成了重大影响。垃圾评论检测（Spam Review Detection）目标就是根据评论内容和评论发布者的行为等信息识别虚假信息。不同于其他垃圾信息识别，虚假评论仅从内容层面非常难进行判断，需要结合用户个体和群体行为才能进行有效的判别。此外，构造虚假评论的标准测试集合也是非常困难的工作，这也使得算法之间不能有效地进行对比。

情感分析相关研究内容还有很多，受篇幅限制本书就不在详述，对情感分析有更多兴趣的读者可以参考刘兵教授关于情感分析任务的专著《Sentiment Analysis: mining sentiments, opinions, and emotions》^[10]。

9.2 篇章级情感分析

篇章级情感分析（Document-level Sentiment Analysis），也称为文档级情感分析，主要任务包括篇章级主客观分类、篇章级极性分类、篇章级情绪分类以及篇章级情感强度判断等情感分类任务^[20-22]。相关分类的目标可以根据本章情感分析概述部分所介绍的观点模型和情绪模型决定。但是，无论分类目标如何，篇章级情感分析任务的可以定义为，对于给定的篇章 d ，预测其所对应的情感标签 y 。如果采用连续评分方法，可以采用回归方法，在损失函数中考虑预测标签之间差距的不同。

篇章级情感分析将整篇文档看做一个整体，通常假设整个文档仅对一个实体进行评论，并不对文档中具体的实体或属性进行细粒度分析。这种假设显然与实际情况并不总是相符的。这不仅制约了篇章级情感分析任务的应用场景，也对模型提出了很大的挑战。篇章中不仅包含主观性的评论内容，也包含客观性的事实描述。比如，电影评论中很可能包含一些对电影情节的描述内容。此外，篇章中除了包含对主要实体或主题的整体评论描述外，也很可能包含对其属性的评价以及对其他相关事物的评价。这些问题都对篇章级情感分析模型造成了很大的挑战，如何在没有句子级标注的情况下区分主观性句子还是客观性句子，如何区分篇章所讨论的主要实体与次要实体，都是篇章级情感分析算法需要解决的问题。

本节主要介绍两类方法：基于支持向量机的篇章级情感分析和基于层次结构的篇章级情感分析。最后，将介绍了一些常见的篇章级情感分析数据集。

9.2.1 基于支持向量机的篇章级情感分析

如前所述，篇章级情感分析通常转换为文本分类问题，因此绝大多数有监督机器学习方法都可以应用于该任务。基于特征的统计机器学习方法，包括朴素贝叶斯、最大熵、支持向量机（Support Vector Machine, SVM）等也都适用于该任务。本节以支持向量机为例，介绍如何利用基于特征的统计学习方法进行篇章级情感分析。

对于给定数据集 $\mathbb{D} = \{(d_1, y_1), (d_2, y_2), \dots, (d_m, y_m)\}$, $y_i \in \{-1, +1\}$ 。对于情感极性分析，我

们可以定义 $y_i = +1$ 表示褒义, $y_i = -1$ 表示贬义。针对主客观分析, $y_i = +1$ 可以定义为表示主观, $y_i = -1$ 表示客观。与其他基于传统机器学习算法的方法一样, 这里也需要首先将文档 d_i 转换为特征表示。利用 SVM 对文本进行分类, 需要将文本转换为向量表示。最基本的方法是采用词袋 (bag-of-words) 特征。给定词典 \mathbb{V} , 该词典中包含了 $|\mathbb{V}|$ 个单词, 那么可以用长度为 $|\mathbb{V}|$ 的向量表示文本。向量的每维表示在文本中某个单词是否出现, 或者采用 TF-IDF 等方法计算得到的该词的权重。Pang 等人在文献 [4] 采用了这种方式分类电影评论取得不错的效果。

在此基础上, 可以将更多与情感分类相关的特征加入向量表示, 下面列出了一些应用于情感分析任务的常见特征:

- 词性: 每一个词的词性特征。形容词和副词是观点和情感的主要承载词, 因此可以通过词性信息的引入使得模型可以利用该特征给相关词语更高的权重。
- 情感词和情感短语: 在语言中表达了积极或者消极情感的词语。例如: 好、很棒等是褒义词, 坏、糟糕等是贬义词。可以利用情感词典将此类单词识别后加入特征。
- 观点的规则: 文本结构和语言成分可以表示隐含情感和观点。可以通过人工设计的规则抽取此类特征加入向量表示中。
- 情感转置词: 可以反转文本中情感倾向的词语或短语。比如否定词可以把正面的情感倾向改变为负面的情感倾向。这类词语可以单独加入提取后加入到向量表示中。
- 句法分析树: 通过句法分析获得的句法分析树或者子树。可以通过将句法路径、句法树片段使用类似 n-gram 的方法转换为向量加入向量表示。此外, SVM 中树核 (Tree Kernel) 方法也可以直接在树结构上进行计算。Wu 等人还提出了针对倾向性分析的树核方法^[23]。

利用上述方法将文档 d_i 转换为对应的特征表示 \mathbf{x}_i 后, 可以采用 SVM 方法构建分类算法。SVM 的基本想法是求解能够且几何间隔最大的分离超平面将不同类别的样本分开 (如图9.4所示)。 $w\mathbf{x} + b = 0$ 即为分离超平面, 距离超平面最近的样本点使式 $w\mathbf{x} + b \geq +1$ 或 $w\mathbf{x} + b \leq -1$ 的等号成立, 这些样本点称为“支持向量” (support vector)。两个类别的支持向量到超平面的距离之和为 $\gamma = \frac{2}{\|\mathbf{w}\|}$, 称为“间隔” (margin)。寻找具有“最大间隔” (maximum margin) 的划分超平面, 也就是寻找满足式约束的参数 w 和 b , 使得间隔最大化。这也等同于最小化 $\frac{1}{\gamma} = \frac{1}{2}\|\mathbf{w}\|^2$ 。整体优化目标可以形式化的表示为:

$$\begin{aligned} \min_{w, b} \quad & \frac{1}{2} \|\mathbf{w}\|^2 \\ \text{s.t. } & y_i (\mathbf{w}^T \mathbf{x}_i + b) \geq 1, \quad i = 1, 2, \dots, m \end{aligned} \quad (9.1)$$

SVM 的基本型是一个凸二次规划 (convex quadratic programming) 问题, 能直接用现成的优化计算方法求解。但 SVM 优化一般使用更高效的拉格朗日乘子法解决。利用拉格朗日乘子法得到其

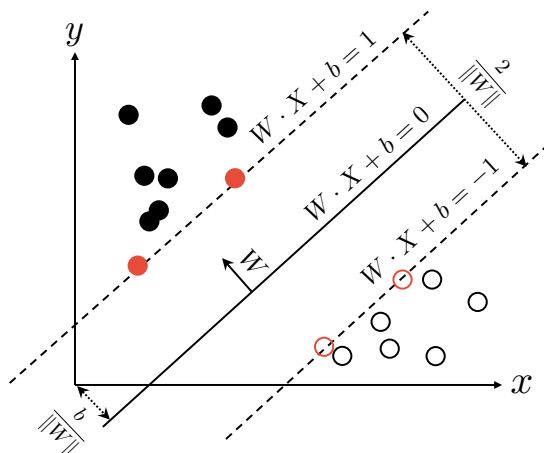


图 9.4 SVM 分类器目标

对偶问题 (dual problem), 将有约束的原始目标函数转换为无约束的新构造的拉格朗日目标函数:

$$\mathcal{L}(\mathbf{w}, b, \boldsymbol{\alpha}) = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^N \alpha_i (y_i (\mathbf{w} \cdot \mathbf{x}_i + b) - 1) \quad (9.2)$$

其中 α_i 为拉格朗日乘子, 且 $\alpha_i \geq 0$ 。

为了得到求解对偶问题的具体形式, 令 $\mathcal{L}(\mathbf{w}, b, \boldsymbol{\alpha})$ 对 \mathbf{w} 和 b 的偏导为 0, 可得:

$$\mathbf{w} = \sum_{i=1}^N \alpha_i y_i \mathbf{x}_i \quad (9.3)$$

$$\sum_{i=1}^N \alpha_i y_i = 0 \quad (9.4)$$

将以上两个等式代入拉格朗日目标函数, 消去 \mathbf{w} 和 b , 可得:

$$\min_{\mathbf{w}, b} \mathcal{L}(\mathbf{w}, b, \boldsymbol{\alpha}) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j) + \sum_{i=1}^N \alpha_i \quad (9.5)$$

求 $\min_{\mathbf{w}, b} \mathcal{L}(\mathbf{w}, b, \boldsymbol{\alpha})$ 对 $\boldsymbol{\alpha}$ 的极大, 即可得到原问题的对偶问题:

$$\begin{aligned}
& \max_{\alpha} \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j \\
& \text{s.t. } \sum_{i=1}^m \alpha_i y_i = 0 \\
& \alpha_i \geq 0, \quad i = 1, 2, \dots, m
\end{aligned} \tag{9.6}$$

利用序列最小优化 (Sequential Minimal Optimiztion, SMO) 算法求解 α , 就可以求解出 w 和 b , 进而求得决策平面。

以上都是基于训练集数据线性可分的假设下进行的, 但是实际情况下几乎不存在完全线性可分的数据, 为了解决这个问题, 还可以进一步引入了“软间隔”的概念, 即允许某些点不满足约束:

$$y_i (\mathbf{w} \cdot \mathbf{x}_i + b) >= 1 \tag{9.7}$$

采用 Hinge 损失, 将原优化问题可以改写为:

$$\begin{aligned}
& \min_{\mathbf{w}, b, \xi_i} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^m \xi_i \\
& \text{s.t. } y_i (\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 - \xi_i \\
& \xi_i \geq 0, \quad i = 1, 2, \dots, N
\end{aligned} \tag{9.8}$$

其中 ξ_i 为“松弛变量”, $\xi_i = \max(0, 1 - y_i (\mathbf{w} \cdot \mathbf{x}_i + b))$ 为一个 Hinge 损失函数, 每一个样本都有一个对应的松弛变量, 表征该样本不满足约束的程度。 $C > 0$ 称为惩罚参数, C 值越大, 对不可分的惩罚越大。于线性可分求解的思路一致, 同样这里先用拉格朗日乘子法得到拉格朗日函数, 再求其对偶问题。

基于上述方法可以完成篇章级主客观分类、篇章级极性分类、篇章级情绪分类, 篇章级情感强度判断等任务^[24, 25]。情感强度判断任务中也可以采用线性回归 (Linear regression)、支持向量回归 (Support vector regression)^[26] 等方法将分数间的差距问题加入目标损失函数。

9.2.2 基于层次结构的篇章级情感分析

篇章具有层次结构, 由单词组成句子, 句子构成篇章。基于这种层次结构, Yang 等人提出了层次注意力网络模型 (Hierarchical attention networks, HAN)^[27] 用于篇章级别情感分析。HAN 的神经网络结构如图9.5所示, 主要包含四个部分: 单词序列编码, 单词级别注意力, 句子序列编码, 句子级别注意力。

单词序列编码层主要用来建模单词序列, 采用了门控循环单元 (Gated Recurrent Unit, GRU) 用来跟踪序列的状态。该模型基于门机制, 不再使用分开的记忆单元。重置门 r_t 和更新门 g_t 用来

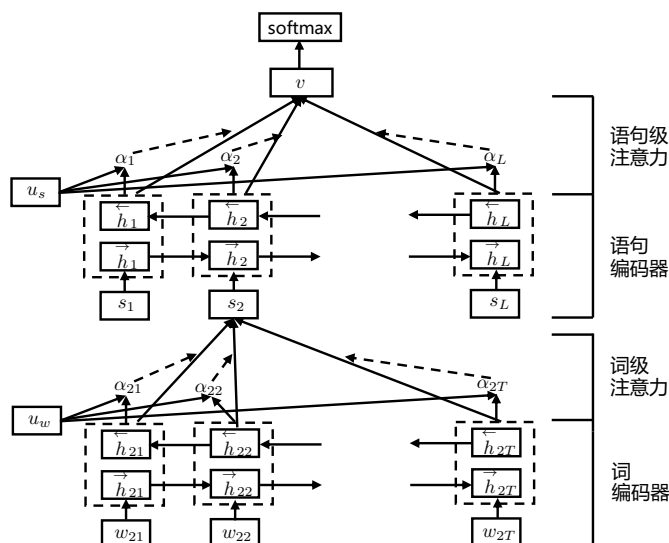


图 9.5 基于层次结构的篇章级情感分析框架----模型结构图重绘

控制状态中信息的更新。更新门 z_t 决定多少信息保留以及多少新的信息加入，重置门 r_t 控制以前状态对于当前状态的影响。在 t 时间，GRU 通过如下公式计算新的状态：

$$\begin{aligned}
 \mathbf{h}_t &= (1 - z_t) \odot \mathbf{h}_{t-1} + z_t \odot \tilde{\mathbf{h}}_t \\
 z_t &= \sigma(\mathbf{W}_z \mathbf{x}_t + \mathbf{U}_z \mathbf{h}_{t-1} + \mathbf{b}_z) \\
 \tilde{\mathbf{h}}_t &= \tanh(\mathbf{W}_h \mathbf{x}_t + \mathbf{r}_t \odot (\mathbf{U}_h \mathbf{h}_{t-1}) + \mathbf{b}_h) \\
 \mathbf{r}_t &= \sigma(\mathbf{W}_r \mathbf{x}_t + \mathbf{U}_r \mathbf{h}_{t-1} + \mathbf{b}_r)
 \end{aligned} \tag{9.9}$$

其中 \mathbf{x}_t 表示时间 t 的序列向量输入， $\mathbf{W}_z, \mathbf{W}_h, \mathbf{W}_r, \mathbf{U}_z, \mathbf{U}_h, \mathbf{U}_r, \mathbf{b}_z, \mathbf{b}_h, \mathbf{b}_r$ 是学习的参数。为此，时间 t 的状态表示 $\mathbf{h}_t = \text{GRU}(\mathbf{x}_t)$ 。

具体地，HAN 中使用双向的 GRU 模型来学习篇章中第 i 个句子 s_i 的第 t 个单词 w_{it} 的表示，分别从单词 w_{i1} 读到 w_{iT} ，从单词 w_{iT} 读到 w_{i1} ，

$$\begin{aligned}
 \vec{\mathbf{h}}_{it} &= \overrightarrow{\text{GRU}}(\mathbf{x}_{it}), t \in [1, T] \\
 \overleftarrow{\mathbf{h}}_{it} &= \overleftarrow{\text{GRU}}(\mathbf{x}_{it}), t \in [T, 1]
 \end{aligned} \tag{9.10}$$

其中 \mathbf{x}_{it} 表示单词 w_{it} 的词向量表示， T 表示句子 s_i 的单词数。最终，正向隐层表示和反向隐层表示拼接 $\mathbf{h}_{it} = [\vec{\mathbf{h}}_{it}; \overleftarrow{\mathbf{h}}_{it}]$ 作为单词 w_{it} 的隐层表示。

并不是句子中所有的单词都对句子的表示同等重要。为此，HAN 中提出使用单词级别的注意

力机制来捕获句子中比较重要的单词，得到加权求和后的句子表示 s_i ：

$$\begin{aligned} \mathbf{u}_{it} &= \tanh(\mathbf{W}_w \mathbf{h}_{it} + \mathbf{b}_w) \\ \alpha_{it} &= \frac{\exp(\mathbf{u}_{it}^\top \mathbf{u}_w)}{\sum_t \exp(\mathbf{u}_{it}^\top \mathbf{u}_w)} \\ \mathbf{s}_i &= \sum_t \alpha_{it} \mathbf{h}_{it} \end{aligned} \quad (9.11)$$

其中 $\mathbf{W}_w, \mathbf{b}_w, \mathbf{u}_w$ 为可以学习的参数。

句子序列编码层和单词序列编码一样，也使用双向的 GRU 模型来编码句子序列，

$$\begin{aligned} \vec{\mathbf{h}}_i &= \overrightarrow{\text{GRU}}(\mathbf{s}_i), i \in [1, |d|] \\ \overleftarrow{\mathbf{h}}_i &= \overleftarrow{\text{GRU}}(\mathbf{s}_i), t \in [|d|, 1] \end{aligned} \quad (9.12)$$

其中 $|d|$ 表示篇章中句子的个数。同样的，通过对于正向和反向句子的拼接，得到最后的句子隐层表示 $\mathbf{h}_i = [\vec{\mathbf{h}}_i; \overleftarrow{\mathbf{h}}_i]$ 。

同样的，在篇章中并不是每一个句子都同等的重要。和单词级别注意力类似，使用句子级别的注意力来捕获篇章中比较重要的句子，得到加权求和后的篇章表示 \mathbf{v} ，

$$\begin{aligned} \mathbf{u}_i &= \tanh(\mathbf{W}_s \mathbf{h}_i + \mathbf{b}_s) \\ \alpha_i &= \frac{\exp(\mathbf{u}_i^\top \mathbf{u}_s)}{\sum_i \exp(\mathbf{u}_i^\top \mathbf{u}_s)} \\ \mathbf{v} &= \sum_i \alpha_i \mathbf{h}_i \end{aligned} \quad (9.13)$$

其中 $\mathbf{W}_s, \mathbf{b}_s, \mathbf{u}_s$ 为可学习参数。最后，通过句子级注意力得到的篇章表示 \mathbf{v} 可以被输入到一个线性模型和 Softmax 函数中用于情感分类预测。

9.2.3 篇章级情感分析语料库

篇章级情感分析通常也依赖大规模的标注语料对模型进行训练和评测。伴随着篇章级情感分析任务的发展，大量篇章级情感分析语料库也相应的提出。本节将介绍几种常见的包含情感标签的篇章级文情感分析语料库。如表9.2所示，主要包含 4 个英文数据集和 2 个中文数据集。

1. 大型电影评论数据集

大型电影评论数据集 (Large Movie Review Dataset) ^[28] 是由斯坦福人工智能实验室 (Stanford Artificial Intelligence Laboratory, SAIL) 于 2011 年推出的一个电影评论的英文数据集，该数据集的训练集和测试集分别包含 25,000 条标注过的电影评论。每一个评论被标记为负面和正面两种类

表 9.2 常见文档级情感分析语料库汇总

语料库名称	训练集合	验证集	测试集	总共	类别	语言
大型电影评论数据集	25,000	-	25,000	50,000	2	英文
IMDB	108,535	13,567	13,567	135,669	10	英文
Yelp-酒店	20,975	6,993	6,993	34,961	5	英文
Yelp-餐厅	106,943	35,648	35,648	178,239	5	英文
Amazon	59,399	11,880	11,880	83,159	5	英文
中文情感语料	-	-	-	1201	2	中文
点评	503,330	83,889	83,889	671,108	5	中文

型。数据集中还额外提供了 50,000 条未标注数据。

2. IMDB

IMDB^[29] 是由卡内基梅隆大学机器学习实验室(Machine Learning Department, Carnegie Mellon University) 和新加坡管理大学于 2014 年联合发布的一个电影评论英文数据集。由于该数据集是篇章级别的数据集, 每一个样本平均包含 393.8 个单词, 14.4 个句子。每一个评论文档被标注 1 到 10 的一个情感极性。该数据集一共包含 135669 个样本, 通过 8:1:1 比例随机划分为训练数据、验证数据和测试数据。

3. Yelp

Yelp^[30] 是由科罗拉多大学波德分校于 2019 年构建的英文数据集数据集。该数据集包含酒店和餐厅两个领域的评论数据。其中酒店领域包含 20,975 条训练样本、6,993 验证样本和 6,993 条测试样本; 餐厅领域包含 106,943 条训练样本、35,648 条验证样本和 35,648 条测试样本。该数据集中每一条评论包含了一个 1-5 的情感打分。对于打分小于 3 的属于消极, 等于 3 的为中性, 而大于 3 的积极。

4. Amazon

Amazon^[30] 也是是由科罗拉多大学波德分校于 2019 年构建的英文数据集数据集, 包含亚马逊上关于音乐的评论数据。该数据规模较大, 一共包含 135,669 条样本, 其中训练样集、验证集和测试集分别包含 108,535、13,567 和 13,567 条样本。和 Yelp 数据类似, 每一个评论都有一个 5 分制的打分, 表示用户的喜欢程度, 值越大对应的情感越积极。

5. 中文情感语料

中文情感语料 (Chinese sentiment corpus, ChnSentiCorp)^[31] 是由中国科学院智能软件实验室(Intelligent Software Department) 于 2008 年构建的篇章级情感分析数据集。该数据集一共包含 1021 个文档, 其中教育相关文档 507 个, 电影相关文档 266 个和房子相关 248 个。每一个文档被分为积极和消极两种。

6. 点评

点评 (Dianping) ^[32] 是一个美团大众点评于 2021 年公布的大型评论数据集。该数据集包含了属性级情感分析和评论预测两个部分。对于评分预测任务, 每一个评论包含一个 1 到 5 的打分, 表示用户对于商品的满意程度。该数据集规模较大, 一共包含大约 67 万个样本, 其中训练样本 503,330 个, 验证样本 93,889 个以及测试样本 93,889 个。

9.3 句子级情感分析

句子级情感分析 (Sentence-level Sentiment Analysis) 主要任务包括句子级主客观分类、句子级极性分类、句子级情绪分类以及句子级情感强度判断等情感分类任务^[33, 34]。给定一个句子 s , 句子级情感分析旨在预测句子对应的情感标签 y 。与篇章级情感分析类似, y 一般指积极、消极或者中性标签。也可以采用连续分数的形式对句子的情感程度进行评分。与篇章级情感分析任务类似, 句子级情感分析虽然包含很多任务, 但是所采用的方法非常类似。

相较于篇章级情感分析, 句子级情感分析则更精细化一些, 将每个句子看做一个整体。通常情况下一个句子级仅对一个实体进行评论。相较于篇章级情感分析中需要对主次实体进行区分的难题, 由于该现象在单个句子中出现次数较少, 因此句子级情感分析中该问题造成影响相对并不严重。但是, 句子级情感分析还要面临新的难题, 主要包含以下几个方面: 1) 句子长度短; 2) 隐式情感表达占比高; 3) 条件、对比等复杂情况多。句子中包含的单词数量少很多, 还包含反讽、双重否定等复杂语言现象, 这就要求模型必须具备充分利用句子中有效的语义信息。隐式观点表达相较于包含明确情感词的显式观点表达的识别难度大很多, 很多情况下还需要基于尝试才能对情感极性进行判断, 而隐式观点在句子级情感分析中的占比很高, 根据在 SemEval-2014 语料集占比达到 30% 左右^[35], 这也对句子级情感分析带来了很大的困难。此外, 句子级情感分析主要关注一般性的分类问题, 但是对于条件、对比等复杂观点句型缺乏处理能力。

本节主要介绍句子级情感分析中一些常见方法, 包括基于无监督方法的句子级情感分析、基于递归神经网络的句子级情感分析、基于预训练的句子级情感分析。最后, 对句子级情感分析的常用的数据集进行说明。

9.3.1 基于词典的句子级情感分析

基于词典的情感分类方法^[36, 37] 是情感分析算法中经典的算法类型之一, 由于该类算法不需要或者仅需要非常少的训练语料, 因此也具有很广泛的应用场景。SO-CAL (Semantic Orientation CALCulator) ^[36, 37] 是 Taboada 等人 2011 年提出的利用情感词典对文本的情感极性进行分析的方法, 其核心是基于对文档中的每个单词或短语根据词典和规则进行分析, 在此基础上综合得到文档情感倾向性极性和强度。SO-CAL 采用了情感字典、情感强化和情感否定等三个方面构建了词典和规则。

情感字典包含形容词（形容词短语）、名词、动词和副词。SO-CAL 所采用的情感字典包含 2,252 个形容词，1,142 个名词，903 个动词和 745 个副词。对于表达正面情感的词和短语给了一个正的值，而对于表达负面的词给一个负的值。具体地，每一个情感词被赋予 -5（极度否定）到 +5（极度肯定）的一个值。

例如：monstrosity -5	hate (noun and verb) -4
disgust -3	sham -3
fabricate -2	delay (noun and verb) -1
determination 1	inspire 2
inspiration 2	endear 3
relish (verb) 4	masterpiece 5

情感强化归纳了情感加强词（比如：非常、很）会增强相邻情感词的语义强度，以及情感弱化词（比如稍微、有点）会减弱相邻情感词的语义强度。文本的情感得分需要结合情感加强词来进行计算。对于加强词和减弱词，给一个或正或负的权重百分比。例如，slightly 是 -50，somewhat 是 -30，pretty 是 -10，really 是 +15，very 是 +25 等。如果 excellent 的情感值是 5，则 very excellent 的情感值为 $5 * (100\% + 25\%)$ 。除了副词和形容词外，还有其他的词性的加强词和减弱词，比如数量词、全部字母大写、感叹号标记、以及语篇连接词 but。

情感否定包含会反转情感词对应的情感极性词汇。在情感分析任务中具有重要作用。情感否定词包括 not、none、nobody、never、nothing 以及其他的单词，如 without 或者 lack。然而，简单的反转情感值会存在一定问题。比如把 excellent 的情感值 +5 通过 not 反转到 -5，并不合理。not excellent 往往比 not good 更加偏向于正向一些。这里采用与原有词语相反方向的固定值相加计算的方式。

例如：这个服务并不是特别好。

特别好：+5，并不：-4，句子整体：+5-4=1

我觉得这个菜的味道这并不差

差：-3，并不：+4，句子整体：-3+4=1

最终，对文本中所有的情感表达值进行求和，若求和为正，则该文本被判定为积极情感，若求和为负，则文本表达一种负面的情感，求和为零，则表示中性情感。例如句子：“这个服务并不是特别好（+5-4=1），但是也并不差（-3+4=-1）”，求和为 0，整体表达的中性的情感。

9.3.2 基于递归神经张量网络的句子级情感分析

句子级情感分析主要是依据句子的语义内容对情感进行分类。递归神经网络 (Recursive Neural Network, RNN) 可以有效利用句法结构，递归的计算句子和短语组合同向量表示，因此可以使用递归神经网络进行句子级情感分析。

每个单词使用 d 维向量进行表示，词向量中的数值使用均匀分布进行随机初始化，所有的词典 \mathbb{V} 的词向量堆叠成一个词嵌入矩阵 $\mathbf{L} \in \mathbb{R}^{d \times |\mathbb{V}|}$ ，与模型一起训练。针对短语“not very good”，其

句法结构如图9.6所示。根据所对应的句法结构，可以按照如下公式，自底向上递归计算每个中间节点的向量表示，计算完两个子节点之后再计算其父结点：

$$p_1 = f\left(\mathbf{W} \begin{bmatrix} b \\ c \end{bmatrix}\right), p_2 = f\left(\mathbf{W} \begin{bmatrix} a \\ p_1 \end{bmatrix}\right) \quad (9.14)$$

其中， $f = \tanh$ 是一个标准的非线性激活函数， $\mathbf{W} \in \mathbb{R}^{d \times 2d}$ 是可以学习的参数。基于得到的节点的向量表示，可以使用 softmax 函数构建分类目标学习函数 $g = \text{softmax}(W_s p_1)$ 。

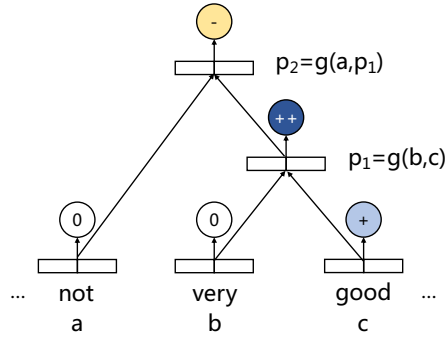


图 9.6 递归神经网络结构图

在此基础上，Socher 等人提出了基于递归神经张量网络（Recursive Neural Tensor Network, RNTN）^[38] 的方法用于句子级情感分析。RNTN 在原有的单词向量表示基础上，增加了基于张量的组合函数用于描述单词之间的组合关系。定义张量积的输出 $\mathbf{h} \in \mathbb{R}^d$ ，通过以下计算得到：

$$\mathbf{h} = \begin{bmatrix} \mathbf{b} \\ \mathbf{c} \end{bmatrix}^T \mathbf{V}^{[1:d]} \begin{bmatrix} \mathbf{b} \\ \mathbf{c} \end{bmatrix} \quad (9.15)$$

其中 $\mathbf{V}^{[1:d]} \in \mathbb{R}^{2d \times 2d \times d}$ 是定义了多个双线性形式，每个 $\mathbf{V}^{[i]}$ 是其中一个切片，使用 $\mathbf{V}^{[i]}$ 计算得到 h_i ：

$$h_i = \begin{bmatrix} \mathbf{b} \\ \mathbf{c} \end{bmatrix}^T \mathbf{V}^{[i]} \begin{bmatrix} \mathbf{b} \\ \mathbf{c} \end{bmatrix} \quad (9.16)$$

计算过程以 $d = 2$ 为例，如图9.7所示。虚线框表示张量的一个切片，用于捕捉到子结点对父节点一种类型的影响。

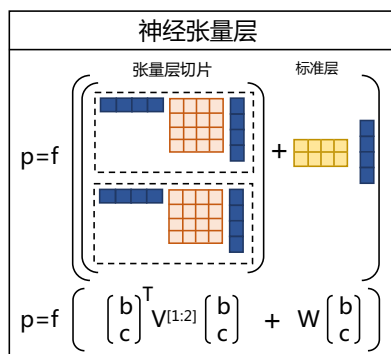


图 9.7 递归神经张量网络单层结构图

在此基础上，图9.6中每个节点的向量表示使用如下公式计算得到：

$$p_1 = f \left(\begin{bmatrix} b \\ c \end{bmatrix}^T V^{[1:d]} \begin{bmatrix} b \\ c \end{bmatrix} + W \begin{bmatrix} b \\ c \end{bmatrix} \right) \quad (9.17)$$

$$p_2 = f \left(\begin{bmatrix} a \\ p_1 \end{bmatrix}^T V^{[1:d]} \begin{bmatrix} a \\ p_1 \end{bmatrix} + W \begin{bmatrix} a \\ p_1 \end{bmatrix} \right) \quad (9.18)$$

根据节点的向量表示 p_1, p_2 ，仍然可以使用 $g = \text{softmax}(W_s p_1)$ 构建分类函数获得句子级情感分析结果。

9.3.3 基于情感知识增强预训练的句子级情感分析

通用预训练模型在绝大部分自然语言处理任务上都取得了很好的效果，基于知识增强的预训练方法进一步提升了预训练模型的在一些自然语言处理任务上的效果。2020年Tian等人提出了情感知识增强的预训练模型SKEP (Sentiment Knowledge Enhanced Pre-training for Sentiment Analysis) 算法^[39]应用于句子级情感分析。SSKEP算法模型框架如图9.8所示。

与BERT^[40]采用的基于掩码语言模型类似，SKEP方法也希望通过掩盖识别输入序列中的情感信息的方法进行模型预训练。但是大规模的情感信息不能通过人工标注完成，因此SKEP方法中提出了使用基于点互信息 (Pointwise Mutual Information, PMI) 的无监督方法识别情感信息^[41]，根据词语的共现信息识别出情感词、情感词的极性和属性词-情感词对等情感信息。掩盖的步骤分为如下三步：

- (1) 掩盖属性词-情感词对：在句子中随机选择最多两对属性词-情感词对掩盖。
- (2) 掩盖情感词：在句子中随机选择不超过当前句子总词数 10% 的情感词进行掩盖。
- (3) 掩盖通用字：如果情感词所占的词比例没有 10%，随机选择单词补充达到总 10% 的掩盖比例。

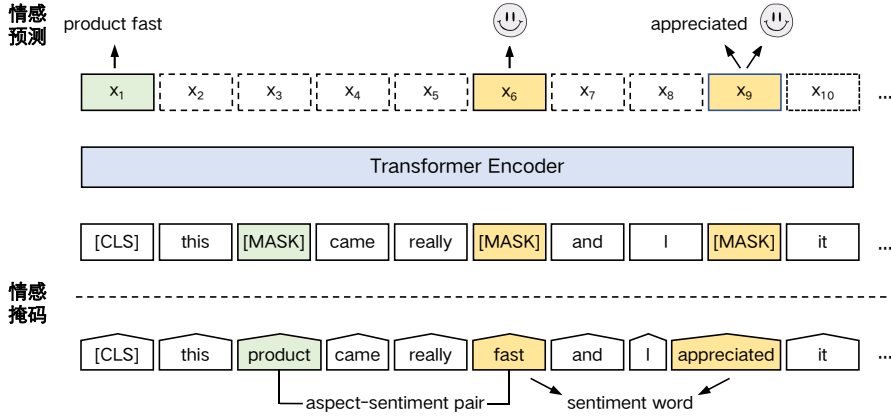


图 9.8 基于情感知识增强预训练的句子级情感分析框架

SKEP 预训练采用的损失函数 L 由三个部分组成：情感词预测目标 (\mathcal{L}_{sw})、情感词极性预测目标 (\mathcal{L}_{wp}) 和属性词-情感词对预测目标 (\mathcal{L}_{ap})。

$$\mathcal{L} = \mathcal{L}_{sw} + \mathcal{L}_{wp} + \mathcal{L}_{ap} \quad (9.19)$$

情感词预测损失 \mathcal{L}_{sw} 定义为：

$$\begin{aligned} \hat{\mathbf{y}}_i &= \text{softmax}(\tilde{\mathbf{x}}_i \mathbf{W} + \mathbf{b}) \\ \mathcal{L}_{sw} &= - \sum_{i=1}^{i=n} m_i \times \mathbf{y}_i \log \hat{\mathbf{y}}_i \end{aligned} \quad (9.20)$$

其中， $\tilde{\mathbf{x}}_i$ 是编码器的输出向量， $\hat{\mathbf{y}}_i$ 是 $\tilde{\mathbf{x}}_i$ 经过输出层后，再经过 softmax 的得到的概率分布。在得到每个位置的预测结果后，并不会计算每个词的损失，而只会计算情感词所在位置的损失，非情感词的位置不会参与计算。 m_i 用于筛选哪些词是情感词。

情感极性词预测损失 \mathcal{L}_{wp} 的计算方式和 \mathcal{L}_{sw} 类似，区别在于 \mathcal{L}_{sp} 计算的是词的损失， \mathcal{L}_{wp} 计算的是极性的损失。属性词-情感词对预测损失 \mathcal{L}_{ap} 则定义为：

$$\begin{aligned} \hat{\mathbf{y}}_a &= \text{sigmoid}(\tilde{\mathbf{x}}_1 \mathbf{W}_{ap} + \mathbf{b}_{ap}) \\ \mathcal{L}_{ap} &= - \sum_{a=1}^{a=A} \mathbf{y}_a \log \hat{\mathbf{y}}_a \end{aligned} \quad (9.21)$$

其中， \mathbf{x}_1 是 [CLS] 位置的输出向量， \mathbf{y}_a 是一个属性词-情感词对， $\tilde{\mathbf{y}}_a$ 是 \mathbf{y}_a 的概率评估值，需要注意的是这里，提前构建了一个个属性词-情感词对的字典库，即每一属性词-情感词对都有一个相应的编号表示。

最终模型预训练完成后，应用于句子级情感分析时，使用 [CLS] 位置输出代表句子整体表示，在编码器上增加分类层，利用句子级情感分析语料进行模型精调。

9.3.4 句子级情感分析语料库

大规模标注语料对于句子级情感分析模型的训练具有重要的作用。本节主要介绍三个比较常用的句子级情感分析语料，包括两个英语数据集，三个中文数据集，如表9.3所示。

表 9.3 常用的句子级情感分析语料库

	训练集合	验证集	测试集	合计	类别	语言
斯坦福情感树库	8,544	1,101	2,210		5	英文
GoEmotions	43,410	5,427	5,426		27	英文
中文情感树库	10,627	665	2,258		2	中文
酒店评论（HR）	-	-	-	24,348	2	中文
冰原历险记三评论（IAR）	-	-	-	11,081	2	中文

1. 斯坦福情感树库

斯坦福情感树库 (Stanford Sentiment Treebank, SST) [38] 数据集由斯坦福大学自然语言处理组于 2013 年发布，包含 8544 条训练数据，1101 条验证数据，2210 条测试数据。其中每个句子分析树的节点均有细粒度的情感注解。句子和短语总有 239232 条，情感值对应类别：[0, 0.2], (0.2, 0.4], (0.4, 0.6], (0.6, 0.8], (0.8, 1.0] 分别对应五分类情感。

2. GoEmotions

GoEmotions[42] 数据集是由谷歌于 2020 年提出的情绪标注数据。该数据集是由 58000 条评论组成的人工注释数据集，这些评论来源于流行英语论坛 Reddit 下的各不同板块，并被标记为 27 个情绪类别。GoEmotions 分类系统是迄今为止最庞大的情绪细化英语数据集，包含详细注释。在设计它时，同时考虑了心理学和数据的适用性。该分类系统包括 12 种积极的、11 种消极的、4 种模糊的以及 1 种“中立”的情绪类别，因此它可以广泛适用于需要对情绪表达进行细致区分的对话理解任务。

3. 中文情感树库

中文情感树库 (Chinese Sentiment Treebank) [43] 数据集是中国科学院自动化研究所于 2014 年构建的数据集。该数据集基于 2270 电影的豆瓣评论，每一个句子都被打上 0 到 4 这样 5 种打分，0 表示非常消息，4 表示非常积极。该数据集过滤了打分为 2 的中性样本，得到 11439 个积极样本和 2111 个消极样本。最终，这些样本被分为训练集（10627 条），验证集（665 条）和测试集（2258 条）。

4. 酒店评论和冰原历险记三评论

酒店评论 (Hotel review, HR) 和冰原历险记三评论 (Ice Age III review, IAR) ^[44] 是由深圳大学于 2015 年提出的中文句子级情感分析数据集。其中 HR 数据集是从携程网站爬取的酒店评论数据, 包含 13,446 条积极和 10,902 消极的评论。IAR 数据集是从豆瓣网站爬取的关于冰原历险记三评论数据, 包含 11,081 条积极和 8,869 条消息的评论。

9.4 属性级情感分析

属性是指被观点表述的对象。例如, “这款手机的电池容量非常好”。这里“电池容量”就是属性。属性级情感分析 (Aspect-level Sentiment Analysis, ABSA) 目标包含属性级情感分类任务和情感信息抽取任务两大类。属性级情感分类任务包括属性级主客观分类、属性级极性分类、属性级情感强度判断。与篇章级和句子级情感分类不同, 属性级情感分类任务的输入不仅包含文本内容 d , 还有目标属性 a , 输出则为文本内容中关于目标属性的观点词 o 和情感标签 y 。情感信息抽取任务目标则是抽取文本的中表达情感的核心要素, 包括评价词、评价对象、观点持有者、评价搭配等。

相较于篇章级和句子的情感分析, 属性级情感分析粒度更小, 因此也称为细粒度情感分析。属性级情感分类任务有效解决了篇章级和句子级情感分类中只能对整体观点进行分析所存在的问题。在实际评论中, 虽然整个篇章或句子对某个实体给出了正面评价, 但不代表对每个属性都给出正面评价, 反之亦然。属性级情感分析可以有效解决上述问题, 在属性层面分别给出细粒度的识别结果。由于属性级情感分析任务更加精细, 因此难度相较于句子级和篇章级情感分析也更大。属性级情感分析不仅要面临句子长度短、隐式情感表达占比高等句子级情感分析所面临的问题。还要处理句子中包含于给定属性相关的内容, 也包含与给定属性无关的内容, 在无相关标注的情况下, 模型也需要具备区分句子中不同部分与给定属性是否相关的能力。

本节主要介绍属性级情感分析所涉及的情感信息抽取和属性级情感分类的常见算法, 包括基于句法规则和基于序列标注的情感信息抽取、基于注意力交互和预训练等方法的属性级情感分类算法。最后, 对属性级情感分析的常用的数据集进行说明。

9.4.1 情感信息抽取

情感信息抽取主要包含属性抽取、观点抽取和情感预测三个子任务。具体地, 给定一个文本序列 $s = \{w_1, \dots, w_{|s|}\}$, 该任务旨在抽取文本中包含的所有的属性、观点、极性三元组 $T = \{(a_i, o_i, p_i)\}_{i=1}^{|T|}$, 其中 $|T|$ 表示样本中三元组的个数, a_i, o_i 分别是第 i 个属性和观点, 是文本序列 X 的一个子串, p_i 是该属性对应的情感极性。

1. 基于句法规则的情感信息抽取

传统基于无监督的情感信息抽取模型主要基于句法信息来抽取文本的属性词和观点词。一般属性词多为名词，而观点词多为形容词，且观点词往往用来修饰属性词。举个例子，对于句子“华为手机拍出来的照片很好看!~”。形容词“好看”直接通过修饰语依赖于名词“照片”。假设观点字典里面只有一个单词“好看”。通过句法修饰关系，可以通过“好看”（观点词）识别出“照片”（属性）。

文献[45]提出通过句法规则来解决观点字典扩展和观点扩展。通过使用依存句法分析器去扩展观点字典和挖掘属性，使得观点词和属性建立联系。使用双向循环的方法来使用信息在观点词和属性之间不断传播。这种方法的好处就是只需要一个初始的观点字典即可。观点字典就是包含许多的观点单词，例如 good、excellent、poor 和 bad。但是使用观点字典的缺陷就在于不可能去囊括所有的意见和领域。并且一个词可能在这个领域是积极的，在另一个领域可能就是中性的。基于句法规则的情感信息抽取方法从已知和被提取的（在之间的迭代）的观点词和属性中，通过识别语义上的联系，迭代提取观点词和属性。这种使得信息在观点词和属性之间来回流动的方法称为双向传播（double propagation）。接下来将从关系识别以及观点字典扩展和属性扩展两个方面来进行算法介绍。

关系识别是指识别观点词和属性（OT-Rel）、属性和属性（TT-Rel）、观点词和观点词（OO-Rel）之间的关系，是观点字典扩展的关键所在。定义了两类单词之间的关系：

- 1) 直接依赖：A 和 B 直接关联（图9.9(a)）、A 和 B 通过 H 直接关联（图9.9(b)）。
- 2) 间接依赖：A 通过 H1 依赖于 B（图9.9(c)）、A 和 B 分别通过 H1 和 H2 依赖于 H（图9.9(d)）。

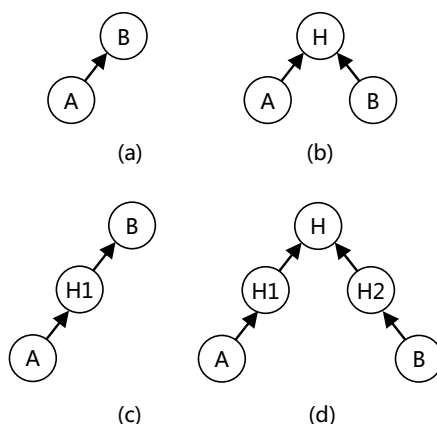


图 9.9 基于规则的属性和观点词抽取

直接依赖和间接依赖都紧紧考虑了句法树的拓扑结构，为了考虑词性，这里还引入了序列标注任务约束。一般观点词多为形容词、属性一般为单个名词或名词短语。具体地，在文献[45]中使

用斯坦福的 POS 工具来抽取词性标注, 根据词性分析结果, 可能的观点词词性为 JJ (Adjectives), JJR (Comparative Adjectives) 和 JJS (Superlative Adjectives)。可能的属性词词性为 NN (Singular Nouns) 和 NNS (Plural Nouns)。描述观点词和属性之间关系的依存关系包括 mod (单词与其直接修饰词关系)、pnmod (名词后修饰语)、subj (动词主语)、s (表层主语)、obj (动词宾语)、obj2 (双及物动词的第二个宾语) 和 desc (描述)。OT-Rel、OO-Rel 或者 TT-Rel 可以形式化定义为一个四元组 $(POS(w_i), DT, R, POS(w_j))$, 其中 $POS(w_i)$ 表示单词 w_i 的词性, DT 表示依存类型 (例如 DD 或者 IDD), R 表示句法关系。

观点字典扩展和属性扩展是指基于预先定义好的规则对观点字典和属性词进行扩张, 主要包含三个部分: 基于关系的传播规则定义、基于规则的传播算法、观点词情感极性预测。

1) 基于关系的传播规则: 主要包含对于进行传播的过程中基于观点词抽取属性、基于额外属性抽取属性、基于额外属性抽取观点词、基于给定和额外的观点词抽取观点词四个子任务的规则 (表9.4)。在表中, o/t 表示抽取的观点词/属性。O/T 表示给定或者抽取的观点词/属性集合。 H 表示任何单词。POS(O/T) 和 O/T-Dep 分别表示单词 O/T 的 POS 信息和依存关系。JJ 和 NN 是潜在的观点词和属性的 POS 标签集合, 其中 JJ 包含 JJ、JJR 和 JJS, NN 包含 NN 和 NNS。MR 是描述观点词和属性关系的依存关系集合, 包含 mod、pnmod、subj、s、obj、obj2 和 desc。CONJ 仅仅包含 conj。箭头表示依存关系。例如, $O \rightarrow O\text{-Dep} \rightarrow T$ 表示 O 通过句法关系 O-Dep 依赖于 T 。“==”表示相同或者相等 (这里相等特指 mod 和 pnmod 一样, s 或者 subj 和 obj 一样)。具体地, 我们利用 $R1_i$ 基于观点词 O 抽取属性 t , $R2_i$ 基于属性 T 抽取观点词 o , $R3_i$ 基于抽取的属性 T_i 抽取属性 t , $R4_i$ 基于已知的观点词 O_i 抽取观点词 o 。以 $R1_1$ 为例, 给定观点词 O , POS 标签为 NN 且满足关系 O-Dep 被抽取为属性词。例如, 对于句子 “The phone has a good screen”, 当我们知道 good 为一个观点词时, 它通过 mod 依赖于 screen, 而 mod 包含在 MR 中且 screen 为 NN, $R1_1$ 会抽取 screen 为属性。

2) 基于规则的传播算法: 输入为观点字典 O 和关于商品的评论数据 R , 其主要思想为: 首先通过初始的观点字典识别句子中的观点词, 然后通过句法关系, 进一步识别出其他观点词或者属性, 然后将他们加入到字典。再不断迭代上面的过程, 直到没有新的观点词和属性能够被识别出来。例如, 对于句子 “Canon G3 takes great pictures, The picture is amazing, You may have to get more storage to store high quality pictures and recorded movies, and The software is amazing”。形容词 great 直接通过 mod 关系依赖于名词 pictures, 定义为 OT-Rel 四元组 (JJ, DD, mod, NN)。假设观点字典里面有一个单词 “great”。根据 great, 通过句法关系, 可以利用 $R1_1$ 可以识别出 picture (属性)。然后通过 picture, 再通过句法关系, 可以在第二句话中利用 $R2_2$ 识别出 amazing。通过 picture 还可以基于 $R3_1$ 识别出 movies。通过 amazing 我们又可以基于 $R1_2$ 识别出 software。

3) 观点词情感极性预测: 通过规则来预测属性观点对应的情感极性。关于观点词和属性, 可以观察到两个现象: 第一, 在一个评论中, 对于相同的属性, 情感极性一般是相同的; 第二, 在一个领域库中, 相同观点词具有相同极性。基于观察到的上述现象, 可以建立下面三条判断情感

表 9.4 属性和观点词抽取规则（来源：文献 [45]）

RuleID	Observations	Output	Example
R1 ₁	$O \rightarrow O\text{-Dep} \rightarrow T$ s.t. $O \in \{O\}$, $O\text{-Dep} \in \{MR\}$, $POS(T) \in \{NN\}$	$t = T$	The phone has a <u>good</u> “screen”. (good \rightarrow mod \rightarrow screen)
R2 ₂	$O \rightarrow O\text{-Dep} \rightarrow H \leftarrow T\text{-Dep}$ $\leftarrow T$ s.t. $O \in \{O\}$, $O/T\text{-Dep} \in \{MR\}$, $POS(T) \in \{NN\}$	$t = T$	“iPod” is the <u>best</u> mp3 player. (best \rightarrow mod \rightarrow player \leftarrow subj \leftarrow iPod)
R2 ₁	$O \rightarrow O\text{-Dep} \rightarrow T$ s.t. $T \in \{T\}$, $O\text{-Dep} \in \{MR\}$, $POS(O) \in \{JJ\}$	$o = O$	same as R1 ₁ with screen as the known word and good as the extracted word
R2 ₂	$O \rightarrow O\text{-Dep} \rightarrow H \leftarrow T\text{-Dep}$ $\leftarrow T$ s.t. $T \in \{T\}$, $O/T\text{-Dep} \in \{MR\}$, $POS(O) \in \{JJ\}$	$o = O$	same as R1 ₂ with iPod as the known word and best as the extract word
R3 ₁	$T_{i(j)} \rightarrow T_{i(j)}\text{-Dep} \rightarrow T_{j(i)}$ s.t. $T_{j(i)} \in \{T\}$, $T_{i(j)}\text{-Dep} \in \{CONJ\}$, $POS(T_{i(j)}) \in \{NN\}$	$t = T_{i(j)}$	Does the player play dvd with <u>audio</u> and “video”? (video \rightarrow conj \rightarrow audio)
R3 ₂	$T_i \rightarrow T_i\text{-Dep} \rightarrow H \leftarrow T_j\text{-Dep} \leftarrow T_j$ s.t. $T_i \in \{T\}$, $T_i\text{-Dep} == T_j\text{-Dep}$, $POS(T_j) \in \{NN\}$	$t = T_j$	Canon “G3” has a great <u>lens</u> . (lens \rightarrow obj \rightarrow has \leftarrow subj \leftarrow G3)
R4 ₁	$O_{i(j)} \rightarrow O_{i(j)}\text{-Dep} \rightarrow O_{j(i)}$ s.t. $O_{j(i)} \in \{O\}$, $O_{i(j)}\text{-Dep} \in \{CONJ\}$, $POS(O_{i(j)}) \in \{JJ\}$	$o = O_{i(j)}$	The camera is <u>amazing</u> and “easy” to use. (easy \rightarrow conj \rightarrow amazing)
R4 ₂	$O_i \rightarrow O_i\text{-Dep} \rightarrow H \leftarrow O_j\text{-Dep} \leftarrow O_j$ s.t. $O_i \in \{O\}$, $O_i\text{-Dep} == O_j\text{-Dep}$, $POS(O_j) \in \{JJ\}$	$o = O_j$	If you want to buy a <u>sexy</u> , “cool”, accessory -available mp3 player, you can choose iPod. (sexy \rightarrow mod \rightarrow player \leftarrow mod \leftarrow cool)

极性的规则：1) 对于由已知属性提取的观点词和由已知观点词提取的属性，赋予它们与已知相同的极性。例如，如果 A 是一个观点词（属性），B 是一个属性（观点词），且 A 通过 B 抽取得到，A 会被赋予 B 一样的情感极性。2) 对于由已知观点词提取的观点词和由已知属性提取的属性，赋予它们与已知相同的极性。除非句子中出现相反的话语。例如，对于单词 A 和 B 都是属性（或者观点词）而 A 通过 B 抽取得到，如果没有相反词在 A 和 B 之间，则 A 被赋予 B 一样的情感极性，反之就是相反的极性。3) 对于一个通过一些属性从其他评论抽取的新观点词，利用整个评论极性来进行预测。评论的极性通过包含的已知的观点词的极性求和得到。如果最终的值大于 0，则为积极，否则为消极。

2. 基于序列标注的情感信息抽取

属性词抽取（Aspect Term Extraction, ATE）是属性级情感分析的一个重要子任务，旨在抽取句子中包含的属性词。具体地，给定一个包含 $|s|$ 个单词的序列 $s = \{w_1, \dots, w_{|s|}\}$ ，ATE 任务目标是

抽取出其中属性词集合 a_1, a_2, \dots, a_n 。该任务通常转换为词级别序列标注问题，通过属性词抽取标签序列 $Y = \{y_1, \dots, y_{|s|}\}$ ，其中 y_i 来自预定义的标签集合 $\mathcal{Y} = \{B, I, O\}$ ，进行属性词抽取。

文献 [46] 针对该任务，提出一个基于序列标注的情感信息抽取模型 ATE-THASTN，该方法同时考虑了抽取的历史和观点信息。ATE-THASTN 算法的神经网络结构如图9.10所示。主要包含两个重要成分：属性历史注意力（Truncated History-Attention, THA）和观点选择网络（Selective Transformation Network, STN）来建模属性历史和观点摘要。历史感知的属性表示和观点摘要拼接在一起作为特征用于当前步的属性预测。

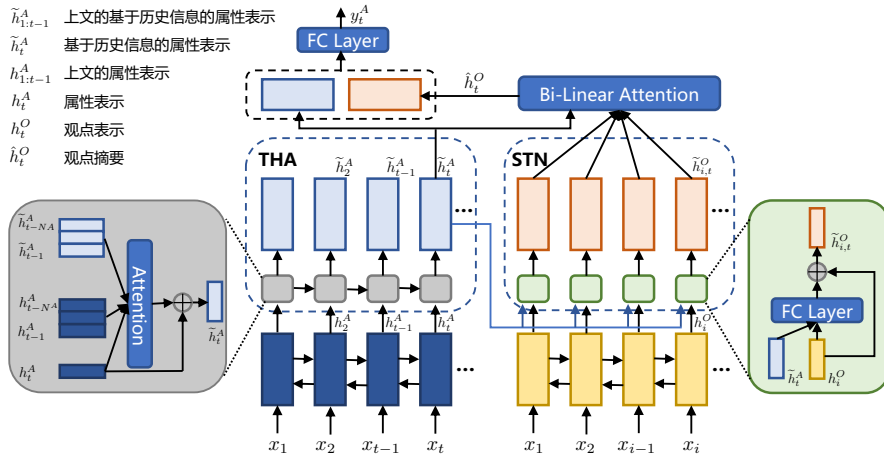


图 9.10 ATE-THASTN 属性抽取神经网络结构^[46]

记忆构建旨在使用两个 LSTM 模型来建模词级别的上下文表示分别用于 ATE 任务和辅助观点词发现任务的序列标注。LSTM^T 表示一个 LSTM 单元，其中 $T \in \{A, O\}$ 是任务指示器。本节中，在没有特定说明的情况下，带有下标 A 和 O 的符号分别表明 ATE 任务和观点发现任务。使用一个双向的 LSTM 模型来生成初始的单词级别表示 h_t^T ，

$$h_t^T = \left[\overrightarrow{LSTM}^T(x_t); \overleftarrow{LSTM}^T(x_t) \right], t \in [1, |s|] \quad (9.22)$$

其中 x_t 表示单词 w_t 的词向量表示。

属性历史建模旨在建模已经预测和当前预测属性之间的关系，建模历史的属性信息。通过考虑 B-I-O 定义减少模型预测当前标签的错误空间，同时提高在一个并列结构的多个属性的预测准确率。一个属性历史注意力模块 (THA) 被提出用于建模属性和属性之间的关系。特别地，THA 缓存最近的 N^A 隐层状态。在当前预测 t 步时刻，THA 计算每一个缓存状态 $h_i^A (i \in [t - N^A, t - 1])$

归一化后的重要度值 s_i^t :

$$\begin{aligned} a_i^t &= \mathbf{v}^\top \tanh \left(\mathbf{W}_1 h_i^A + \mathbf{W}_2 h_t^A + \mathbf{W}_3 \tilde{h}_i^A \right) \\ s_i^t &= \text{Softmax} \left(a_i^t \right) \end{aligned} \quad (9.23)$$

\tilde{h}_i^A 表示前面历史感知的属性表示。然后属性历史 \hat{h}_t^A 可以计算为:

$$\hat{h}_t^A = \sum_{i=t-N^A}^{t-1} s_i^t \times \tilde{h}_i^A \quad (9.24)$$

为了利用已经抽取的属性, 将隐层属性表示和蒸馏后的属性历史合并来生成当前预测的特征。采用和残差块类似的结构来计算第 t 步历史感知的属性表示 \tilde{h}_t^A ,

$$\tilde{h}_t^A = h_t^A + \text{ReLU} \left(\hat{h}_t^A \right) \quad (9.25)$$

其中 ReLU 表示 relu 激活函数。

观点摘要建模旨在利用观点选择网络 (STN) 来选择和属性相关的观点信息从而抑制可能的噪音。在处理全局的观点之前插入 STN, 从而抓取对于给定属性候选更加重要的特征。特别地, 给定当前属性特征 \tilde{h}_t^A , STN 首先计算一个新的观点表示 $\hat{h}_{i,t}^O$,

$$\hat{h}_{i,t}^O = h_i^O + \text{ReLU} \left(\mathbf{W}_4 \tilde{h}_t^A + \mathbf{W}_5 h_i^O \right) \quad (9.26)$$

其中 \mathbf{W}_4 、 \mathbf{W}_5 是可以学习的参数。属性的特征 \tilde{h}_t^A 在这里类似一个过滤器来保留重要的观点特征。

对于蒸馏后的全局观点摘要, 使用一个双向线性项来计算 \tilde{h}_t^A 和 $\hat{h}_{i,t}^O$ 之间的相关值:

$$w_{i,t} = \text{Softmax} \left(\tanh \left(\tilde{h}_t^A \mathbf{W}_{bi} \hat{h}_{i,t}^O + \mathbf{b}_{bi} \right) \right) \quad (9.27)$$

其中 \mathbf{W}_{bi} 和 \mathbf{b}_{bi} 是双向线性注意力的参数。第 t 步, 增强的观点表示通过观点表示的加权求和得到:

$$\hat{h}_t^O = \sum_{i=1}^T w_{i,t} \times \hat{h}_{i,t}^O \quad (9.28)$$

最后, 将观点摘要 \hat{h}_t^O 和历史感知的属性表示 \tilde{h}_t^A 进行拼接, 输入到全连接层用于属性预测:

$$\begin{aligned} f_t^A &= \left[\tilde{h}_t^A : \hat{h}_t^O \right] \\ P \left(y_t^A \mid x_t \right) &= \text{Softmax} \left(\mathbf{W}_f^A f_t^A + \mathbf{b}_f^A \right) \end{aligned} \quad (9.29)$$

该框架利用多任务进行训练，包括属性预测和观点预测。初始化的词级别表示 h_i^O 被作为特征进行观点预测：

$$P(y_i^O | x_i) = \text{Softmax}(\mathbf{W}_f^O h_i^O + \mathbf{b}_f^O) \quad (9.30)$$

词级别的预测的分布 $P(y_t^T | x_t) (\mathcal{T} \in \{A, O\})$ 和真实分布 $P(y_t^{T,g} | x_t)$ 的交叉熵错误作为损失函数：

$$\mathcal{L}_T = -\frac{1}{T} \sum_1^T P(y_t^{T,g} | x_t) \cdot \log[P(y_t^T | x_t)] \quad (9.31)$$

属性抽取 \mathcal{L}_A 和观点抽取 \mathcal{L}_O 两个损失函数求和作为最后的损失函数：

$$\mathcal{L} = \mathcal{L}_A + \mathcal{L}_O \quad (9.32)$$

9.4.2 属性级情感分类

与篇章和句子级别的情感分类任务不同，属性级情感分类任务关注句子中评价对象的特定属性的情感，也称细粒度情感分类、属性级情感分类和属性感知的情感分类。一个句子的中可能包含评价对象的多个属性，并且句子中对于不同属性所表达情感可能不一致，甚至完全相反。

例如：我买了一台新相机，照片质量很好，但是电池寿命太短。

该句中关于相机的“图片质量”和“电池寿命”这两个属性的情感极性不同，同时观点词“很好”和“太短”分别用来修饰这两个属性，表达了积极和消极的情感。属性级别情感分析任务目标是预测句子对于给定属性的情感。可以形式化定义为：给定一个包含 $|s|$ 个单词都句子 $s = \{w_1, w_2, \dots, w_{|s|}\}$ 以及一个包含 $|A|$ 个属性的列表 $A = \{a_1, \dots, a_{|A|}\}$ ，其中每一个属性 $a_i = \{w_{i_1}, \dots, w_{i_{|a_i|}}\}$ 是句子 s 的一个子序列。目标是输出属性列表中每个属性在当前文本中的情感极性 $P = \{p_1, \dots, p_{|A|}\}$ 。同时，为了减少误差传播，现有工作联合进行情感信息抽取和属性级情感分类任务。具体地，给定句子 s ，抽取所有的属性词 $A = \{a_1, \dots, a_{|A|}\}$ 和观点词 $O = \{o_1, \dots, o_{|O|}\}$ ，再对属性词 a_i 和观点词 o_i 进行匹配并预测其情感极性 p_i ，得到最后三元组集合 $T = \{(a_i, o_i, p_i)\}_{i=1}^{|T|}$ 。

1. 基于概率混合模型的属性级情感分类

文献[47]提出一个基于混合语言模型的主题-情感模型(Topic-Sentiment Model, TSM)来同时抽取文本中包含的主题和情感。过程形式化地表示为：对于给定的一个文档集合 $\mathcal{C} = \{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_{|\mathcal{C}|}\}$ ，假设语料中包含了 \parallel 个主要主题 $\{\theta_1, \theta_2, \dots, \theta_k\}$ 。每一个主题可以建模为字典中所有单词上的多项式分布(Multinomial Distribution)。在主题-情感模型中，每一个词被分为通用词(例如“这”、“那”、“的”)和主题词两种，而对于一个主题的单词，会进一步分为三个子类别：(1) 带有中性观点的主题词(例如“价格”)；(2) 有关主题积极观点的词(例如“爱”、“喜欢”)；(3) 有关主题消极观点的词(例如“讨厌”、“差”)。该主题-情感模型包含四个多项式分布：(1) θ_B 表示用来抽取通用词的背景主题模型；(2) $\Theta = \{\theta_1, \dots, \theta_k\}$ 表示用来抽取关于 k 个子主题中性描述的 k 个主题模型；(3)

θ_P 表示用来抽取积极观点的积极情感模型；(4) θ_N 表示用来抽取消极观点的消极情感模型。基于概率混合模型的属性级情感分类概率图模型表示如图9.11所示。

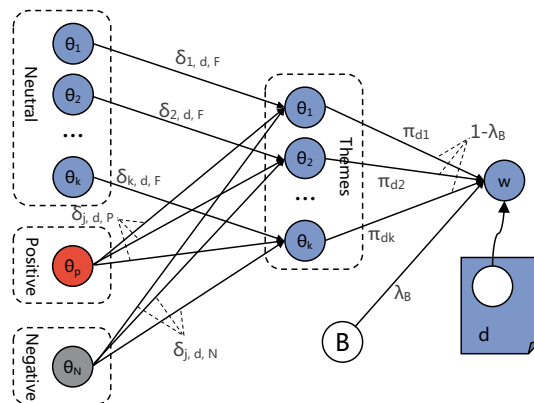


图 9.11 基于概率混合模型的属性级情感分类概率图模型表示^[47]

基于上述概率图模型，具体地抽取流程包括以下四步：1) 判断一个单词是否为通用词。如果是，则该单词通过 θ_B 来采样；2) 如果不是通用词，则判断该词属于 k 个主题词的哪一个；3) 判断主题之后，进一步判断对于主题是中性、积极还是消极的；4) 是的步骤 2 中选中主题为第 j 个主题 θ_j ，基于步骤 3，通过 θ_j ， θ_P 或者 θ_N 来采样单词。

根据 TSM 模型，整个语料 \mathcal{C} 的对数似然为：

$$\log(\mathcal{C}) = \sum_{d \in \mathcal{C}} \sum_{w \in V} c(w : d) \log \left[\lambda_B p(w | B) + (1 - \lambda_B) \sum_{j=1}^k \pi_{dj} \times (\delta_{j,d,F} p(w | \theta_j) + \delta_{j,d,P} p(w | \theta_P) + \delta_{j,d,N} p(w | \theta_N)) \right] \quad (9.33)$$

其中 $c(w : d)$ 是文档 d 中词 w 的个数， λ_B 表示选择 θ_B 的概率，这里作为一个 0 和 1 之间的实验常数。 π_{dj} 表示文档 d 中选择第 j 个主题的概率， $\{\delta_{j,d,F}, \delta_{j,d,P}, \delta_{j,d,N}\}$ 表示文档 d 中主题 j 中性、积极和消极观点的情感覆盖度。

背景模型被定义为：

$$p(w | \theta_B) = \frac{\sum_{d \in \mathcal{C}} c(w, d)}{\sum_{w \in V} \sum_{d \in \mathcal{C}} c(w, d)} \quad (9.34)$$

剩下还需要被估计的参数包括主题模型参数 $\Theta = \{\theta_1, \dots, \theta_k\}$ ，情感模型参数 θ_P 和 θ_N ，文档主题概率 π_{dj} ，每一个文档的情感覆盖度 $\{\delta_{j,d,F}, \delta_{j,d,P}, \delta_{j,d,N}\}$ 。这里将这些参数定义为 Λ 。这些参数可以通过期望最大化 (Expectation-Maximization, EM) 算法来迭代计算最大似然估计，更新方

式如公式9.35:

$$\begin{aligned}
 p(z_{d,w,j,F} = 1) &= \frac{(1 - \lambda_B) \pi_{dj}^{(n)} \delta_{j,d,F}^{(n)} p^{(n)}(w | \theta_j)}{\lambda_B p(w | \theta_B) + (1 - \lambda_B) \sum_{j'=1}^k \pi_{dj'}^{(n)} \left(\delta_{j',d,F}^{(n)} p^{(n)}(w | \theta_{j'}) + \delta_{j',d,P}^{(n)} p^{(n)}(w | \theta_P) + \delta_{j',d,N}^{(n)} p^{(n)}(w | \theta_N) \right)} \\
 p(z_{d,w,j,P} = 1) &= \frac{(1 - \lambda_B) \pi_{dj}^{(n)} \delta_{j,d,P}^{(n)} p^{(n)}(w | \theta_P)}{\lambda_B p(w | \theta_B) + (1 - \lambda_B) \sum_{j'=1}^k \pi_{dj'}^{(n)} \left(\delta_{j',d,F}^{(n)} p^{(n)}(w | \theta_{j'}) + \delta_{j',d,P}^{(n)} p^{(n)}(w | \theta_P) + \delta_{j',d,N}^{(n)} p^{(n)}(w | \theta_N) \right)} \\
 p(z_{d,w,j,N} = 1) &= \frac{(1 - \lambda_B) \pi_{dj}^{(n)} \delta_{j,d,N}^{(n)} p^{(n)}(w | \theta_N)}{\lambda_B p(w | \theta_B) + (1 - \lambda_B) \sum_{j'=1}^k \pi_{dj'}^{(n)} \left(\delta_{j',d,F}^{(n)} p^{(n)}(w | \theta_{j'}) + \delta_{j',d,P}^{(n)} p^{(n)}(w | \theta_P) + \delta_{j',d,N}^{(n)} p^{(n)}(w | \theta_N) \right)} \\
 \pi_{dj}^{(n+1)} &= \frac{\sum_{w \in V} c(w, d) (p(z_{d,w,j,F} = 1) + p(z_{d,w,j,P} = 1) + p(z_{d,w,j,N} = 1))}{\sum_{j'=1}^k \sum_{w \in V} c(w, d) (p(z_{d,w,j',F} = 1) + p(z_{d,w,j',P} = 1) + p(z_{d,w,j',N} = 1))} \\
 \delta_{j,d,F}^{(n+1)} &= \frac{\sum_{w \in V} c(w, d) p(z_{d,w,j,F} = 1)}{\sum_{w \in V} c(w, d) (p(z_{d,w,j,F} = 1) + p(z_{d,w,j,P} = 1) + p(z_{d,w,j,N} = 1))} \\
 \delta_{j,d,P}^{(n+1)} &= \frac{\sum_{w \in V} c(w, d) p(z_{d,w,j,P} = 1)}{\sum_{w \in V} c(w, d) (p(z_{d,w,j,F} = 1) + p(z_{d,w,j,P} = 1) + p(z_{d,w,j,N} = 1))} \\
 \delta_{j,d,N}^{(n+1)} &= \frac{\sum_{w \in V} c(w, d) p(z_{d,w,j,N} = 1)}{\sum_{w \in V} c(w, d) (p(z_{d,w,j,F} = 1) + p(z_{d,w,j,P} = 1) + p(z_{d,w,j,N} = 1))} \\
 p^{(n+1)}(w | \theta_j) &= \frac{\sum_{d \in \mathcal{C}} c(w, d) p(z_{d,w,j,F} = 1)}{\sum_{w' \in V} \sum_{d \in \mathcal{C}} c(w', d) p(z_{d,w',j,F} = 1)} \\
 p^{(n+1)}(w | \theta_P) &= \frac{\sum_{d \in \mathcal{C}} \sum_{j=1}^k c(w, d) p(z_{d,w,j,P} = 1)}{\sum_{w' \in V} \sum_{d \in \mathcal{C}} \sum_{j=1}^k c(w', d) p(z_{d,w',j,P} = 1)} \\
 p^{(n+1)}(w | \theta_N) &= \frac{\sum_{d \in \mathcal{C}} \sum_{j=1}^k c(w, d) p(z_{d,w,j,N} = 1)}{\sum_{w' \in V} \sum_{d \in \mathcal{C}} \sum_{j=1}^k c(w', d) p(z_{d,w',j,N} = 1)}
 \end{aligned} \tag{9.35}$$

其中 $\{z_{d,w,j,s}\}$ 是隐层变量的集合 ($s \in \{F, P, N\}$), $p(z_{d,w,j,s})$ 表示基于主题/情感模型 w , 文档 d 中单词 w 从第 j 个主题生成的概率。

在实际应用中如果不对模型的增加任何约束, 仅通过 EM 算法优化得到的情感模型会偏向于特定的内容, 主题模型也会因为情感被产生偏差。这主要是因为观点词和主题词经常共现, 很难被 EM 算法分开。这导致情感模型和主题模型独立, 而主题模型应该是中性的。为此, TSM 算法提出先定义模型先验从而约束情感/主题模型学习这个先验, 再利用最大后验概率 (Maximum a Posterior, MAP) 估计将这些先验与数据似然性结合来预估参数。

通过定义模型先验从而约束情感/主题模型学习这个先验。假设 $\bar{\theta}_P$ 和 $\bar{\theta}_N$ 为根据训练语料学习到的积极和消极情感模型。为情感模型 θ_P 和 θ_N 定义了两个共轭狄利克雷先验, $Dir(\{1 + u_P p(w|\bar{\theta}_P)\}_{w \in V})$ 和 $Dir(\{1 + u_N p(w|\bar{\theta}_N)\}_{w \in V})$, 其中参数 u_P 和 u_N 表明对于情感模型先验的自信程度。由于先验是共轭的, 所以 u_P (或者 u_N) 可以解释为“等效样本容量”。也就是说, 当评估情感模型 $p(w|\theta_P)$ (或者 $p(w|\theta_N)$) 时, 添加先验的影响和添加词 w 的 $u_P p(w|\bar{\theta}_P)$ (或者 $u_N p(w|\bar{\theta}_N)$) 伪数目等效。

为此，模型中所有参数的先验假设为：

$$p(\Lambda) \propto p(\theta_P) * p(\theta_N) * \prod_{j=1}^k p(\theta_j) = \prod_{w \in V} p(w | \theta_P)^{\mu_P p(w | \bar{\theta}_P)} \prod_{w \in V} p(w | \theta_N)^{\mu_N p(w | \bar{\theta}_N)} \prod_{j=1}^k \prod_{w \in V} p(w | \theta_j)^{\mu_j p(w | \bar{\theta}_j)} \quad (9.36)$$

其中当没有对与 θ_j 的先验知识时， $u_j = 0$ 。

最大后验概率估计则是基于以上定义的先验知识，使用 MAP 估计 $\hat{\Lambda} = \arg \max_{\Lambda} p(C | \Lambda) p(\Lambda)$ 。为了结合先验给定的伪数目，TSM 对 EM 算法中的 M 步进行了重写。新的 M 步更新如下：

$$\begin{aligned} p^{(n+1)}(w | \theta_P) &= \frac{\mu_P p(w | \bar{\theta}_P) + \sum_{d \in \mathcal{C}} \sum_{j=1}^k c(w, d) p(z_{d,w,j,P} = 1)}{\mu_P + \sum_{w' \in V} \sum_{d \in \mathcal{C}} \sum_{j=1}^k c(w', d) p(z_{d,w',j,P} = 1)} \\ p^{(n+1)}(w | \theta_N) &= \frac{\mu_N p(w | \bar{\theta}_N) + \sum_{d \in \mathcal{C}} \sum_{j=1}^k c(w, d) p(z_{d,w,j,N} = 1)}{\mu_N + \sum_{w' \in V} \sum_{d \in \mathcal{C}} \sum_{j=1}^k c(w', d) p(z_{d,w',j,N} = 1)} \\ p^{(n+1)}(w | \theta_j) &= \frac{\mu_j p(w | \bar{\theta}_j) + \sum_{d \in \mathcal{C}} c(w, d) p(z_{d,w,j,F} = 1)}{\mu_j + \sum_{w' \in V} \sum_{d \in \mathcal{C}} c(w', d) p(z_{d,w',j,F} = 1)} \end{aligned} \quad (9.37)$$

参数 u_P , u_N , u_j 可以是一个常数，也可以是通过正则估计来设定。

训练后模型可以用于以下应用：

- (1) 句子主题排序：给定句子集合和一个主题 j ，通过对于主题 j 的得分的进行句子排序：

$$\text{Score}_j(s) = -D(\theta_j \| \theta_s) = - \sum_{w \in V} p(w | \theta_j) \log \frac{p(w | \theta_j)}{p(w | \theta_s)} \quad (9.38)$$

其中 θ_s 是一个句子 s 的平滑语言模型。

- (2) 句子情感分类：给定主题为 j 的句子 s ，预测其积极、消极或者中性情感：

$$\arg \max_x -D(\theta_s \| \theta_x) = \arg \max_x - \sum_{w \in V} p(w | \theta_s) \log \frac{p(w | \theta_s)}{p(w | \theta_x)} \quad (9.39)$$

其中 $x \in \{j, P, N\}$ ， θ_s 是 s 的语言模型。

- (3) 预测文档或者主题的整体观点：给定一个文档 d 和一个主题 j ，主题 j 在文档 d 的整体情感分布为情感覆盖度 $\{\delta_{j,d,F}, \delta_{j,d,P}, \delta_{j,d,N}\}$ 。对于主题 j 的整体情感强度为：

$$S(j, P) = \frac{\sum_{d \in \mathcal{C}} \pi_{dj} \delta_{j,d,P}}{\sum_{d \in \mathcal{C}} \pi_{dj}} \quad (9.40)$$

2. 基于注意力交互的属性级情感分类

由于一个句子中可能对评价对象的多个属性都进行了评论，因此如何建模属性和上下文之间交互关系，从而获取和属性相关的上下文，对于属性级情感分类任务有非常重要的作用。针对该问题，文献 [48] 提出了多粒度注意力网络 MGAN (Multi-grained Attention Network) 算法，同时考虑粗粒度和细粒度的注意力交互机制，用于建模属性和上下文之间的交互关系，其模型结构如图9.12所示。MGAN 主要包含输入嵌入层、上下文建模层、多粒度注意力层和输出层四个部分。

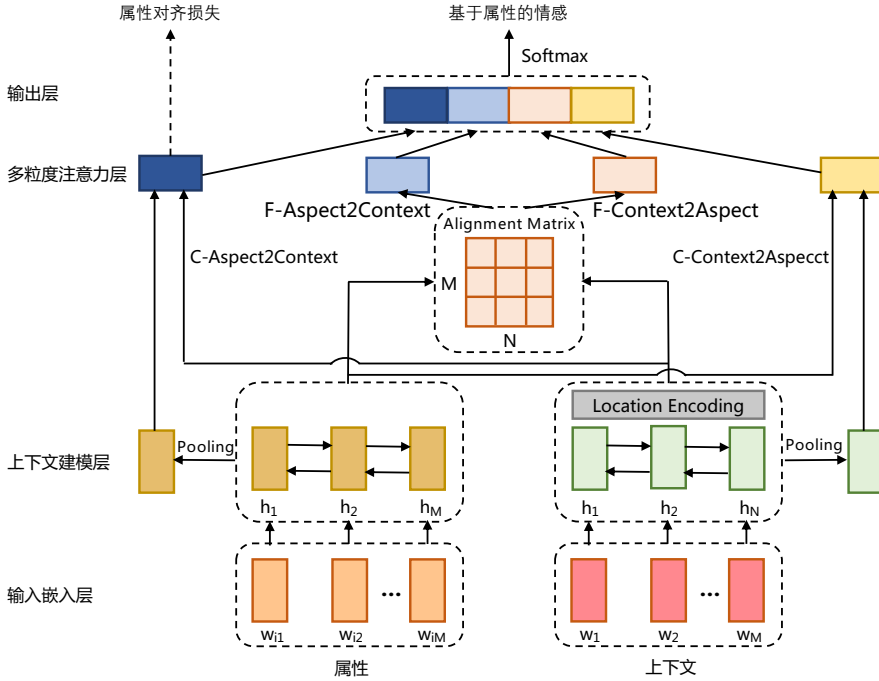


图 9.12 基于注意力交互的属性级情感分类模型神经网络结构

输入嵌入层将每一个单词映射到高维空间。MGAN 使用了预训练的单词向量 Glove 来获得固定的每一个单词的词嵌入表示。嵌入矩阵定义为 $\mathbf{L} \in \mathbb{R}^{d_v \times |V|}$ ，其中 d_v 是单词向量的维度， $|V|$ 是词表大小。

上下文建模层使用双向的 LSTM 模型来建模句子的时间序列关系。给定上下文句子 s 和对应属性 a_j 的单词词嵌入，使用两个单独的 BiLSTM 进行建模，句子表示为 $\mathbf{H} \in \mathbb{R}^{2d \times s}$ 和属性表示为 $\mathbf{Q} \in \mathbb{R}^{2d \times |a_j|}$ 。同时，由于距离属性越近的单词对于属性的影响更大，为此位置编码机制将该

引入融合考虑。对于一个离属性具体为 l 的上下文单词 w_j , 其权重定义如下:

$$\alpha_i = 1 - \frac{l}{|s| - |a_i| + 1} \quad (9.41)$$

特别地, 对于属性内部的词, 权重设定为 0。由此, 可以得到上下文表示 $\mathbf{H} = [\mathbf{H}_1 * \alpha_1, \dots, \mathbf{H}_{|s|} * \alpha_{|s|}]$ 。

多粒度注意力层从粗粒度和细粒度两个角度进行属性和句子之间的交互。为了关联属性与上下文信息, MGAN 算法提出了细粒度注意力机制, 通过扑捉单词级别的交互来估计每一个属性和上下文单词影响。细粒度和粗粒度的注意力向量被拼接来获得最后的表示。由于属性之间的关系, 也可以提供额外有价值的信息。MGAN 算法还设计了属性对齐损失函数用来加强属性和同一个上下文以及不同情感极性的注意力差别。

粗粒度注意力用于建模属性和上下文之间的交互, 使用属性的平均值来计算上下文单词的注意力权重。这里使用了双向注意力机制, 分别叫做 C-Aspect2Context 和 C-Context2Aspect。

- (1) C-Aspect2Context 根据属性向量的平均值学习属性对于上下文单词的权重。这里对属性上下文输出 \mathbf{Q} 使用平均池化来生成平均属性向量 $\mathbf{Q}_{avg} \in \mathbb{R}^{2d}$ 。对于每一个上下文中的单词向量 \mathbf{H}_i , 通过如下公式计算注意力值 \mathbf{a}_i^{ca} :

$$\begin{aligned} s_{ca}(\mathbf{Q}_{avg}, \mathbf{H}_i) &= \mathbf{Q}_{avg} * \mathbf{W}_{ca} * \mathbf{H}_i \\ \mathbf{a}_i^{ca} &= \frac{\exp(s_{ca}(\mathbf{Q}_{avg}, \mathbf{H}_i))}{\sum_{k=1}^N \exp(s_{ca}(\mathbf{Q}_{avg}, \mathbf{H}_k))} \end{aligned} \quad (9.42)$$

其中值函数 s_{ca} 计算属性对于上下文单词的重要性权重。 $\mathbf{W}_{ca} \in \mathbb{R}^{2d \times 2d}$ 是注意力权重矩阵。上下文输出的加权求和表示 $\mathbf{m}^{ca} \in \mathbb{R}^{2d}$ 可通过如下公式计算得到:

$$\mathbf{m}^{ca} = \sum_{i=1}^N \mathbf{a}_i^{ca} \cdot \mathbf{H}_i \quad (9.43)$$

- (2) C-Context2Aspect 学习上下文对于属性单词的权重。与 C-Aspect2Context 类似, 使用平均池化来获得平均上下文向量 \mathbf{H}_{avg} 来计算每一个属性中单词 w_i 的权重。最终属性向量的加权求和表示 \mathbf{bmm}^{cc} 可通过如下公式计算得到:

$$\begin{aligned} s_{cc}(\mathbf{H}_{avg}, \mathbf{Q}_i) &= \mathbf{H}_{avg} * \mathbf{W}_{cc} * \mathbf{Q}_i \\ \mathbf{a}_i^{cc} &= \frac{\exp(s_{cc}(\mathbf{H}_{avg}, \mathbf{Q}_i))}{\sum_{k=1}^M \exp(s_{cc}(\mathbf{H}_{avg}, \mathbf{Q}_k))} \\ \mathbf{m}^{cc} &= \sum_{i=1}^M \mathbf{a}_i^{cc} \cdot \mathbf{Q}_i \end{aligned} \quad (9.44)$$

细粒度注意力旨在建模词级别的交互，目标是评估每一个属性单词如何影响其他上下文单词。上下文 H 和属性 Q 之间的对齐矩阵 U 表示第 i 个上下文单词和第 j 个属性单词之间的相似度。该矩阵具体计算方法如下：

$$U_{ij} = W_u ([H_i; Q_j; H_i * Q_j]) \quad (9.45)$$

- (1) F-Aspect2Context 计算对于其中一个属性词最相关的上下文词。对于上下文词的注意力权重 a_i^{fa} 计算方法如下：

$$\begin{aligned} s_i^{fa} &= \max(U_{i,:}) \\ a_i^{fa} &= \frac{\exp(s_i^{fa})}{\sum_{k=1}^N \exp(s_k^{fa})} \end{aligned} \quad (9.46)$$

根据权重可以获得注意力向量 m^{fa} ：

$$m^{fa} = \sum_{i=1}^N a_i^{fa} \cdot H_i \quad (9.47)$$

- (2) F-Context2Aspect 计算对于其中一个上下文词最相关的属性词。基于上下文词的注意力权重 a^{fc} ，得到最后的向量表示 q^{fc} ：

$$\begin{aligned} a_{ij}^{fc} &= \frac{\exp(U_{ij})}{\sum_{k=1}^M \exp(U_{ik})} \\ q_i^{fc} &= \sum_{j=1}^M a_{ij}^{fc} \cdot Q_j \end{aligned} \quad (9.48)$$

最终，对 q^{fc} 进行均值池化来获得聚合后的向量 m^{fc} ：

$$m^{fc} = \text{Pooling} \left([q_1^{fc}, \dots, q_N^{fc}] \right) \quad (9.49)$$

粗粒度和细粒度注意力向量拼接起来作为最后的表示，并将其输入到输出层来预测属性情感极性。

$$\begin{aligned} m &= [m^{ca}; m^{cc}; m^{fa}; m^{fc}] \\ p &= \text{softmax}(W_p * m + b_p) \end{aligned} \quad (9.50)$$

为了能够使得不同属性的注意力的不同，在 C-Aspect2Context 注意力权重上还加入属性对齐损失函数。在该损失函数的约束下，每一个属性会通过和其他相关属性的对比更加关注重要的单词。具体地，对于属性列表 A 中每一个属性对 a_i 和 a_j ，计算粗粒度注意力向量 a_i^{ca} 和 a_j^{ca} 的平方

损失并估计 a_i 和 a_j 的距离 d_{ij} 作为损失权重:

$$d_{ij} = \sigma(\mathbf{W}_d([\mathbf{Q}_i; \mathbf{Q}_j; \mathbf{Q}_i * \mathbf{Q}_j]))$$

$$\mathcal{L}_{align} = - \sum_{i=1}^{M-1} \sum_{j=i+1, y_i \neq y_j}^M \sum_{k=1}^N d_{ij} \cdot (\mathbf{a}_{ik}^{ca} - \mathbf{a}_{jk}^{ca})^2 \quad (9.51)$$

其中 σ 是 sigmoid 函数, y_i 和 y_j 是属性 a_i 和 a_j 的标签, \mathbf{a}_{ik}^{ca} 和 \mathbf{a}_{jk}^{ca} 表示属性 a_i 和 a_j 对于第 k 个上下文单词的注意力权重。

最后的损失函数包含交叉熵损失、属性对齐损失和正则项:

$$\mathcal{L} = - \sum_{i=1}^C y_i \log(p_i) + \beta \mathcal{L}_{align} + \lambda \|\theta\|^2 \quad (9.52)$$

其中 β 和 λ 为需要预先给出的超参数。

3. 基于端到端的联合属性级情感分类

属性级情感分类通常转化为情感要素抽取和分类两个步骤, 但是两个任务级联会导致错误的传播, 从而影响最终结果。为了解决上述问题, 文献 [49] 提出了基于生成框架的情感信息抽取和属性级情感分析统一框架, 其神经网络结构如图9.13所示。该模型将属性级情感分析任务分为抽取和分类两个子任务, 可以分别表示为指针索引和类别索引。将这两个子任务形式化到一个同一个的生成框架下。

使用 a , s , o 分别表示属性词, 情感极性和观点词。上标 s 和 e 分别表示一个词开始和结束索引。例如: o^s 和 a^e 表示观点词 o 的开始索引和属性词 a 的结束索引, s^p 表示情感极性类别的索引, 根据上述标签, 每个子任务对应的目标序列如下:

属性词抽取 (AE) : $Y = [a_1^s, a_1^e, \dots, a_i^s, a_i^e, \dots]$,

观点词抽取 (OE) : $Y = [o_1^s, o_1^e, \dots, o_i^s, o_i^e, \dots]$

属性级情感分类 (AESC) : $Y = [a_1^s, a_1^e, s_1^p, \dots, a_i^s, a_i^e, s_i^p, \dots]$,

二元组抽取: $Y = [a_1^s, a_1^e, o_1^s, o_1^e, \dots, a_i^s, a_i^e, o_i^s, o_i^e, \dots]$,

三元组抽取: $Y = [a_1^s, a_1^e, o_1^s, o_1^e, s_1^p, \dots, a_i^s, a_i^e, o_i^s, o_i^e, s_i^p, \dots]$,

以上的子任务仅仅关注输入句子, 然而属性级情感分类 (ALSC) 和属性词观点词关系抽取 (AOE) 两个子任务依赖于特定的属性 a 。该模型没有将属性词作为输入, 而是放入目标端, 如下所示:

属性级情感分类 (ALSC) : $Y = [\underline{a}^s, \underline{a}^e, s^p]$,

属性词观点词关系抽取 (AOE) : $Y = [\underline{a}^s, \underline{a}^e, o_1^s, o_1^e, \dots, o_i^s, o_i^e, \dots]$,

在推理过程中, 带有下划线的部分是给定的。对于每一个子任务具体的目标序列样例如图9.14所示。

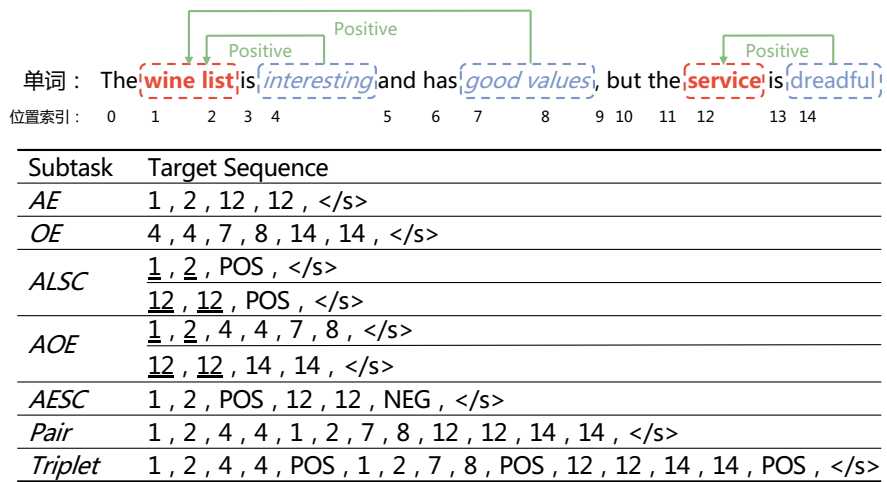
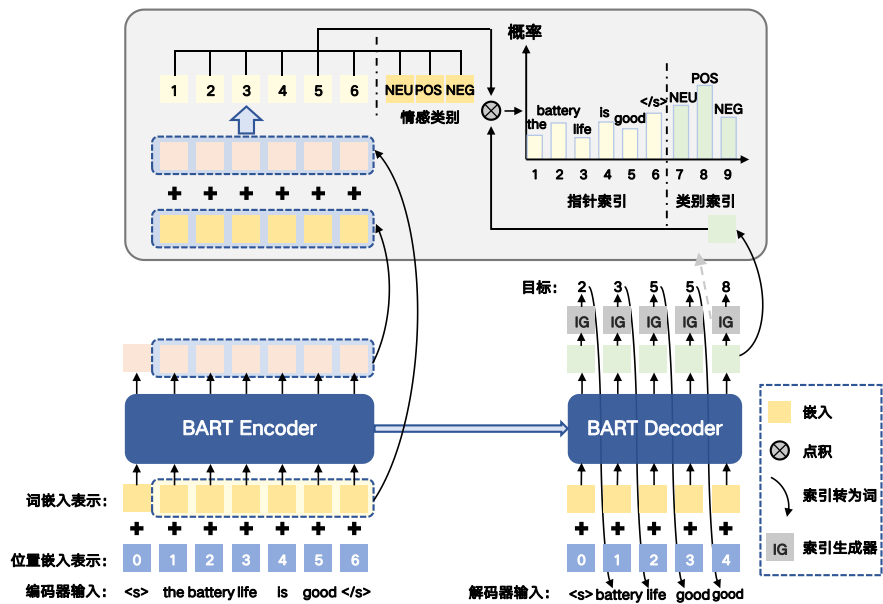


图 9.14 生成目标序列样例

基于上述表示, 所有的子任务都可以形式化定义为以 $X = [x_1, \dots, x_n]$ 为输入, 目标序列 $Y =$

$[y_1, \dots, y_m]$ 为输出, 其中 y_0 是句子开始的标志词。不同的属性级情感分析任务可以定义为:

$$P(Y|X) = \prod_{t=1}^m P(y_t|X, Y_{<t}). \quad (9.53)$$

为了获得每一步索引概率分布 $P_t = P(y_t|X, Y_{<t})$ 使用了编码器和解码器两个模块。

编码器将 X 编码为向量 \mathbf{H}^e 。这里使用 BART 模型。句子的开始词 ($\langle s \rangle$) 和结束词 ($\langle /s \rangle$) 添加到 X 的开始和结束位置。这里为了简单处理, 在公式中忽略了 $\langle s \rangle$ 词。编码器部分如下:

$$\mathbf{H}^e = \text{BARTEncoder}([x_1, \dots, x_n]), \quad (9.54)$$

其中 $\mathbf{H}^e \in \mathbb{R}^{n \times d}$, d 为隐层维度。

解码器把编码器的输出 \mathbf{H}^e 和前面解码器的输出 $Y_{<t}$ 作为输入来计算 P_t 。然而, $Y_{<t}$ 是一个索引序列, 为此, 对于每一个 $Y_{<t}$ 中的 y_t , 需要使用下面的索引到词的模型来做一个转化:

$$\hat{y}_t = \begin{cases} X_{y_t}, & \text{如果 } y_t \text{ 是一个指针索引} \\ C_{y_t - n} = s, & \text{如果 } y_t \text{ 是一个类别索引} \end{cases}$$

其中 $C = [c_1, \dots, c_l]$ 类别词列表。

最后, 使用 BART 解码器来获得最后的隐层表示

$$h_t^d = \text{BARTDecoder}(\mathbf{H}^e; \hat{Y}_{<t}), \quad (9.55)$$

基于 h_t^d , 使用如下公式预测词的概率分布 P_t :

$$\begin{aligned} \mathbf{E}^e &= \text{BARTTokenEmbed}(X), \\ \hat{\mathbf{H}}^e &= \text{MLP}(\mathbf{H}^e), \\ \bar{\mathbf{H}}^e &= \alpha \hat{\mathbf{H}}^e + (1 - \alpha) \mathbf{E}^e, \\ C^d &= \text{BARTTokenEmbed}(C), \\ P_t &= \text{Softmax}([\bar{\mathbf{H}}^e; C^d] h_t^d), \end{aligned} \quad (9.56)$$

在训练的过程中, 使用 Teacher-forcing 来训练模型。具体地, 该策略在训练网络过程中, 每次不使用上一个状态的输出作为下一个状态的输入, 而是直接使用训练数据的标准答案 (ground truth) 的对应上一项作为下一个状态的输入。通过这种方式可以矫正模型的预测, 避免在序列生成的过程中误差进一步放大。负对数似然 (Negative log-likelihood, NLL) 被用来作为损失函数优化模型。更多地, 在推理阶段, 使用束搜索 (beam search) 来获得目标序列 Y 。束搜索有一个超参数

beam size (束宽), 设为 k 。第一个时间步长, 选取当前条件概率最大的 k 个词, 当做候选输出序列的第一个词。之后的每个时间步长, 基于上个步长的输出序列, 挑选出所有组合中条件概率最大的 k 个, 作为该时间步长下的候选输出序列。始终保持 k 个候选, 最后从 k 个候选中挑出最优的。最后, 基于获得目标序列 Y , 使用解码算法将序列转化为词片段和情感极性。即将指针索引根据其中文本中开始和结束位置转化为具体的属性词或者观点词, 将类别索引转化为具体的情感类别。

9.4.3 属性级情感分析语料库

目前使用较多的属性级情感分类数据集是语义评测国际研讨会发布^[50-52], 包括 SemEval 2014^[50], SemEval 2015^[51] 和 SemEval 2016^[52]。同时, 包括 Twitter^[53], Sentihood^[54] 也常用于该任务。常用的属性级情感分析语料库如表9.5所示。

表 9.5 常见的属性级情感分析语料库

	训练集合	验证集	测试集	合计	任务	语言
Restaurant14	1,978	-	600	2,578	分类	英文
Laptop14	1,462	-	411	1,873	分类	英文
Restaurant15	1,120	-	582	1,702	分类	英文
Restaurant16	1,708	-	587	2,295	分类	英文
Twitter	6,248	-	692	6,940	分类	英文
Sentihood	-	-	-	5215	分类	英文
MPQA					抽取	英文
ASAP	36,850	4,490	4,490	46,730	抽取	中文

1. SemEval 2014-2016

SemEval 2014 task4^①^[50] 关注的是基于属性级别情感分析, 该任务的目标是检测给定目标实体的属性并确定每个属性所表达的情感极性。有两个针对笔记本电脑和餐馆的特定领域的数据集, 即 Restaurants14 和 Laptop14。每个句子都被归入句子中讨论的以下五个属性的一个或多个类别: (1) 食物; (2) 服务; (3) 价格; (4) 氛围 (指餐厅的气氛和环境的句子); (5) 轶事/杂事 (不属于上述四个类别的句子)。具体来说, 每个单字或多字的属性词都根据句子中对它所表达的情感而被赋予以下极性之一: (1) 积极; (2) 消极; (3) 中性 (指既非积极也非消极的情绪); (4) 冲突 (意味着既是积极又是消极的情绪)。SemEval-2015 任务 12^②^[51] 是 SemEval-2014 任务 4 的延续。SemEval-2016 任务 5^③^[52] 与 SemEval-2015 任务 12 类似, 该数据集由整个评论组成。此外, 该数据集包含五个领域, 涵盖八种语言。

① <http://alt.qcri.org/semeval2014/task4/>

② <http://alt.qcri.org/semeval2015/task12/>

③ <http://alt.qcri.org/semeval2016/task5/>

Restaurants14 Restaurants14 由从 Ganu 等人 [55] 从餐厅评论/评论中提取的 3000 多条英文句子组成，作为训练数据集。额外评论以相同的方式进行标注作为测试数据集。在去除有冲突的情绪极性或没有属性词的数据后，剩下 1978 个训练样本和 600 个测试样本。该数据集包括对粗略的属性类别、属性词、属性词特定极性和属性类别特定极性的标注。

Laptop14 这个数据集由超过 3000 个从客户笔记本评论中获得的英文句子组成。该数据集的一部分被划分为测试数据。在除去有冲突的情感极性或没有属性词的数据后，剩下 1462 个训练样本和 411 个测试样本。该数据集只包括对句子的属性词及其极性的标注。

Restaurants15 该数据集由 254 条和 96 条餐厅评论组成，分别为训练和测试的属性及其情感极性做了标注。每个评论可能包含多个句子，每个句子包括类别、属性词和属性极性的标注。在去除冲突情绪极性的数据后，有 1120 句用于训练，582 句用于测试。

Restaurants16 该数据集由 350 条餐厅评论组成，其中有用于训练的属性词、属性类别和极性的标注，有 92 条用于测试。在去除有冲突情绪极性的数据后，有 1708 个标注的句子用于训练，587 个用于测试。

2. Twitter

Twitter 数据集^[53]是由北京航空大学、微软亚洲研究院和哈尔滨工业大学于 2014 年联合发布的属性级情感分析数据集。该数据集为手动标注的数据集，用于属性依赖的推特情感分析。这是最大的属性依赖的 twitter 情感分类数据集，它是由人工标注的。训练数据有 6248 条推文，测试数据包括 692 条推文，其情感类别平衡为 25% 负面，50% 中性，25% 正面。

3. Sentihood

SentiHood^[54]是由英国伦敦大学学院、华威大学等公布的是一个城市街区领域中属性级情感分析基准英文数据集。它是基于与伦敦市街区有关的问题，这些问题是通过过滤雅虎答案的问题回答平台的文本而获得的。SentiHood 由 5215 个句子组成，其中 3862 个句子包含一个地点，1353 个句子包含两个地点。在整个数据集中，位置实体名称被 location1 和 location2 所掩盖。

4. MPQA

MPQA^① [56] 是由匹兹堡大学于 2015 年发布的属性级情感分析英文数据集，起包含了标注了观点和其他状态（如情感、信念、情绪和猜测）的新闻文章和其他文本文件。在 MPQA 3.0 中，增加了实体-属性和事件-属性（eTarget）标注。一共有 70 个文档，除了 MPQA 2.0 的 1029 个表达主观元素、1287 个态度和 1213 个属性片段外，还加入了 1366 个 eTargets 到表达主观元素，1608 个 eTargets 到属性片段。

① <http://mpqa.cs.pitt.edu/corpora/>

5. ASAP

ASAP^[57]是由大众点评于2021年发布的属性级情感分析中文语料库。该库由46730条真实世界的用户评论组成,其被随机分成训练集(36,850)、验证集(4,940)和测试集(4,940),包含了18个属性类别。数据集标注了文本中包含的情感以及每个属性相应的情感极性,分为积极、中性和消极。

9.5 延伸阅读

本章介绍了文档级情感分析、句子级情感分析和属性级情感分析三个任务情感分析常见算法。这些方法大多是基于大规模标签数据的有监督机器学习方法,在样本不足情况下表现较差。同时,情感分析任务需要借助大量的外部知识,如句法分析、情感知识库等,如何更好利用这些知识也成为一个重要问题。对于属性级情感分析任务,还存在如何建模属性和上下文之间的交互等问题。近年来,很多研究者从不同方面对上述问题进行探索。

为了解决数据集较小的问题,研究者提出很多更加高效的建模算法,包括递归神经网络^[38, 58–60],卷积神经网络(CNN)^[61, 62],层次结构模型^[27, 62, 63]等。同时,许多研究通过微调的方式将预训练的模型应用于情感分析等下游任务^[40, 64–66]。Song等人^[66–68]集成BERT进行属性的情感分类,并取得了显著的进步,这表明预训练这种迁移学习的有效性。更多地,基于迁移学习情感分析的方法也得到学术研究者的广泛关注,包括任务迁移^[69],基于枢轴的(Pivot-based)跨领域迁移模型^[70, 71],基于对抗的跨领域迁移模型^[72–74]等。

为了更好地利用现有的知识,大量的研究工作对如何结合外部知识用于情感分析任务进行探索。一些工作提出将外部知识添加到预训练BERT中,以增强表示^[75, 76]。Levine等人^[77]引入了SenseBERT来通过预测WordNet中的标记的同义词来提高词汇理解。田等人^[39]和Ke等人^[78]结合外部知识以学习情感信息。此外,也有研究将情感相关概念的常识知识纳入属性级情感分析任务的深度神经网络的端到端训练中,包括循环神经网络模型^[79],图神经网络模型^[80]等。为了考虑依赖关系树,采用递归神经网络^[53, 81]、注意力网络^[82]、依赖关系树结构感知模型^[83, 84]进行建模。

属性级情感分析任务的其中一个挑战是如何获取给定属性的相关信息。传统方法依赖于设计好的特征,导致构建这些模型相当费力^[85]。由于神经网络模型在分布式方面上非常大的优势,目前的研究也大都着重在神经网络模型方面,包括循环神经网络模型^[86],深度记忆网络(Deep memory network)^[87, 88],卷积神经网络模型^[89],门机制^[90],层次网络^[91]等。注意力模型在捕获与给定属性相关的重要部分的方面体现强大能力,包括属性感知的注意力机制^[92, 93],多层次注意力机制^[88],交互式注意力机制^[92]。使用句子中某个属性的位置信息来捕获更准确的特定于属性的信息的想法吸引了研究人员的关注^[83, 84, 88, 94]。

9.6 习题

- (1) 当情感表达为隐式情感，没有情感词时如何训练情感分类器？
- (2) 目前的情感分析系统可解释性和鲁棒性如何？以及如何判断模型的情感可解释性和鲁棒性？
- (3) 情感分析和文本分类有什么区别？属性级情感分析和信息抽取有什么区别？
- (4) 如何进行跨语言情感分析？
- (5) 跨领域情感倾向分析的主要难点有哪些？

参考文献

- [1] WIEBE J M. Learning subjective adjectives from corpora[C]//Proceedings of the 17th National Conference on Artificial Intelligence (AAAI-2000). 2000.
- [2] Das S, Chan M. Extracting market sentiment from stock message boards[J]. Asia Pacific Finance Association, 2001, 2001.
- [3] Tong R M. An operational system for detecting and tracking opinions in on-line discussion[C]//Working Notes of the ACM SIGIR 2001 Workshop on Operational Text Classification: volume 1. 2001.
- [4] Pang B, Lee L, Vaithyanathan S. Thumbs up?: sentiment classification using machine learning techniques[C]//Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10. Association for Computational Linguistics, 2002: 79-86.
- [5] Nasukawa T, Yi J. Sentiment analysis: Capturing favorability using natural language processing[C]//Proceedings of the 2nd international conference on Knowledge capture. 2003: 70-77.
- [6] Dave K, Lawrence S, Pennock D M. Mining the peanut gallery: Opinion extraction and semantic classification of product reviews[C]//Proceedings of the 12th international conference on World Wide Web. 2003: 519-528.
- [7] Wiebe J. Identifying subjective characters in narrative[C]//COLING 1990 Volume 2: Papers presented to the 13th International Conference on Computational Linguistics. 1990.
- [8] Wiebe J M. Tracking point of view in narrative[J]. Computational Linguistics, 1994, 20(2):233-287.
- [9] Wiebe J, Bruce R, O' Hara T P. Development and use of a gold-standard data set for subjectivity classifications[C]//Proceedings of the 37th annual meeting of the Association for Computational Linguistics. 1999: 246-253.
- [10] Liu B. Sentiment analysis: Mining opinions, sentiments, and emotions[M]. Cambridge university press, 2020.

- [11] Liu B. Web data mining: exploring hyperlinks, contents, and usage data: volume 1[M]. Springer, 2011.
- [12] Jindal N, Liu B. Mining comparative sentences and relations[C]//Aaai: volume 22. 2006: 9.
- [13] Parrott W G. Emotions in social psychology: Essential readings[M]. psychology press, 2001.
- [14] Plutchik R. The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice[J]. American scientist, 2001, 89(4): 344-350.
- [15] Mehrabian A, Russell J A. An approach to environmental psychology.[M]. the MIT Press, 1974.
- [16] Russell J A. Evidence of convergent validity on the dimensions of affect.[J]. Journal of personality and social psychology, 1978, 36(10).
- [17] Pang B, Lee L, et al. Opinion mining and sentiment analysis[J]. Foundations and Trends® in Information Retrieval, 2008, 2(1-2):1-135.
- [18] Liu B. Sentiment analysis and opinion mining[J]. Synthesis lectures on human language technologies, 2012, 5(1):1-167.
- [19] Liu B, Hu M, Cheng J. Opinion observer: analyzing and comparing opinions on the web[C]//Proceedings of the 14th international conference on World Wide Web. 2005: 342-351.
- [20] Behdenna S, Barigou F, Belalem G. Document level sentiment analysis: a survey[J]. EAI Endorsed Transactions on Context-aware Systems and Applications, 2018, 4(13).
- [21] Behdenna S, Barigou F, Belalem G. Sentiment analysis at document level[C]//International Conference on Smart Trends for Information Technology and Computer Communications. Springer, 2016: 159-168.
- [22] Tsytarau M, Palpanas T. Survey on mining subjective data on the web[J]. Data Mining and Knowledge Discovery, 2012, 24(3):478-514.
- [23] Wu Y, Zhang Q, Huang X J, et al. Phrase dependency parsing for opinion mining[C]//Proceedings of the 2009 conference on empirical methods in natural language processing. 2009: 1533-1541.
- [24] Moraes R, Valiati J F, Neto W P G. Document-level sentiment classification: An empirical comparison between svm and ann[J]. Expert Systems with Applications, 2013, 40(2):621-633.

- [25] Tripathy A, Anand A, Rath S K. Document-level sentiment classification using hybrid machine learning approach[J]. Knowledge and Information Systems, 2017, 53(3):805-831.
- [26] Drucker H, Burges C J, Kaufman L, et al. Support vector regression machines[J]. Advances in neural information processing systems, 1996, 9.
- [27] Yang Z, Yang D, Dyer C, et al. Hierarchical attention networks for document classification[C]// Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2016: 1480-1489.
- [28] Maas A, Daly R E, Pham P T, et al. Learning word vectors for sentiment analysis[C]//Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies. 2011: 142-150.
- [29] Diao Q, Qiu M, Wu C Y, et al. Jointly modeling aspects, ratings and sentiments for movie recommendation (jmars)[C]//Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. 2014: 193-202.
- [30] Huang X, Paul M J. Neural temporality adaptation for document classification: Diachronic word embeddings and domain adaptation models[C/OL]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence, Italy: Association for Computational Linguistics, 2019: 4113-4123. <https://aclanthology.org/P19-1403>. DOI: 10.18653/v1/P19-1403.
- [31] Tan S, Zhang J. An empirical study of sentiment analysis for chinese documents[J]. Expert Systems with applications, 2008, 34(4):2622-2629.
- [32] Bu J, Ren L, Zheng S, et al. ASAP: A Chinese review dataset towards aspect category sentiment analysis and rating prediction[C/OL]//Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Online: Association for Computational Linguistics, 2021: 2069-2079. <https://www.aclweb.org/anthology/2021.naacl-main.167>.
- [33] Bongirwar V K. A survey on sentence level sentiment analysis[J]. International Journal of Computer Science Trends and Technology (IJCTST), 2015, 3(3):110-113.
- [34] Jagtap V, Pawar K. Analysis of different approaches to sentence-level sentiment classification[J]. International Journal of Scientific Engineering and Technology, 2013, 2(3):164-170.

- [35] Li Z, Zou Y, Zhang C, et al. Learning implicit sentiment in aspect-based sentiment analysis with supervised contrastive pre-training[C]//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. 2021: 246-256.
- [36] Taboada M, Brooke J, Tofiloski M, et al. Lexicon-based methods for sentiment analysis[J]. Computational linguistics, 2011, 37(2):267-307.
- [37] Hu M, Liu B. Mining and summarizing customer reviews[C]//Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. 2004: 168-177.
- [38] Socher R, Perelygin A, Wu J, et al. Recursive deep models for semantic compositionality over a sentiment treebank[C]//Proceedings of the 2013 conference on empirical methods in natural language processing. 2013: 1631-1642.
- [39] Tian H, Gao C, Xiao X, et al. Skep: Sentiment knowledge enhanced pre-training for sentiment analysis[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020: 4067-4076.
- [40] Devlin J, Chang M W, Lee K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). 2019: 4171-4186.
- [41] Turney P. Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews[C]//Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics. 2002: 417-424.
- [42] Demszky D, Movshovitz-Attias D, Ko J, et al. Goemotions: A dataset of fine-grained emotions[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020: 4040-4054.
- [43] Li C, Xu B, Wu G, et al. Recursive deep learning for sentiment analysis over social data[C]//2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT): volume 2. IEEE, 2014: 180-185.
- [44] Fu X, Xu Y. Recursive autoencoder with hownet lexicon for sentence-level sentiment analysis[M]//Proceedings of the ASE BigData & SocialInformatics 2015. 2015: 1-7.
- [45] Qiu G, Liu B, Bu J, et al. Opinion word expansion and target extraction through double propagation [J]. Computational linguistics, 2011, 37(1):9-27.

- [46] Li X, Bing L, Li P, et al. Aspect term extraction with history attention and selective transformation [C/OL]//Lang J. Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, July 13-19, 2018, Stockholm, Sweden. ijcai.org, 2018: 4194-4200. <https://doi.org/10.24963/ijcai.2018/583>.
- [47] Mei Q, Ling X, Wondra M, et al. Topic sentiment mixture: modeling facets and opinions in weblogs [C]//Proceedings of the 16th international conference on World Wide Web. 2007: 171-180.
- [48] Fan F, Feng Y, Zhao D. Multi-grained attention network for aspect-level sentiment classification[C]// Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. 2018: 3433-3442.
- [49] Yan H, Dai J, Ji T, et al. A unified generative framework for aspect-based sentiment analysis[C/OL]// Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). On-line: Association for Computational Linguistics, 2021: 2416-2429. <https://aclanthology.org/2021.acl-long.188>. DOI: 10.18653/v1/2021.acl-long.188.
- [50] Manandhar S. Semeval-2014 task 4: Aspect based sentiment analysis[C]//Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014). 2014.
- [51] Pontiki M, Galanis D, Papageorgiou H, et al. Semeval-2015 task 12: Aspect based sentiment analysis [C]//Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015). 2015: 486-495.
- [52] Pontiki M, Galanis D, Papageorgiou H, et al. Semeval-2016 task 5: Aspect based sentiment analysis [C]//Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016). 2016: 19-30.
- [53] Dong L, Wei F, Tan C, et al. Adaptive recursive neural network for target-dependent twitter sentiment classification[C]//Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers): volume 2. 2014: 49-54.
- [54] Saeidi M, Bouchard G, Liakata M, et al. Sentihood: Targeted aspect based sentiment analysis dataset for urban neighbourhoods[C]//Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. 2016: 1546-1556.
- [55] Ganu G, Elhadad N, Marian A. Beyond the stars: Improving rating predictions using review text content.[C]//WebDB: volume 9. Citeseer, 2009: 1-6.

- [56] Deng L, Wiebe J. Mpqa 3.0: An entity/event-level sentiment corpus[C]//Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2015: 1323-1328.
- [57] Bu J, Ren L, Zheng S, et al. Asap: A chinese review dataset towards aspect category sentiment analysis and rating prediction[C]//Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2021: 2069-2079.
- [58] Socher R, Lin C C, Manning C, et al. Parsing natural scenes and natural language with recursive neural networks[C]//Proceedings of the 28th international conference on machine learning (ICML-11). 2011: 129-136.
- [59] Socher R, Pennington J, Huang E H, et al. Semi-supervised recursive autoencoders for predicting sentiment distributions[C]//Proceedings of the conference on empirical methods in natural language processing. Association for Computational Linguistics, 2011: 151-161.
- [60] Socher R, Huval B, Manning C D, et al. Semantic compositionality through recursive matrix-vector spaces[C]//EMNLP. Association for Computational Linguistics, 2012: 1201-1211.
- [61] Kim Y. Convolutional neural networks for sentence classification[J]. arXiv preprint arXiv:1408.5882, 2014.
- [62] Tang D, Qin B, Liu T. Learning semantic representations of users and products for document level sentiment classification[C]//Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). 2015: 1014-1023.
- [63] Chen H, Sun M, Tu C, et al. Neural sentiment classification with user and product attention[C]//EMNLP. 2016.
- [64] Dodge J, Ilharco G, Schwartz R, et al. Fine-tuning pretrained language models: Weight initializations, data orders, and early stopping[J]. arXiv preprint arXiv:2002.06305, 2020.
- [65] Sun C, Qiu X, Xu Y, et al. How to fine-tune bert for text classification?[C]//China National Conference on Chinese Computational Linguistics. Springer, 2019: 194-206.
- [66] Xu H, Liu B, Shu L, et al. Bert post-training for review reading comprehension and aspect-based sentiment analysis[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics. 2019.

- [67] Song Y, Wang J, Jiang T, et al. Attentional encoder network for targeted sentiment classification[J]. arXiv preprint arXiv:1902.09314, 2019.
- [68] Sun C, Huang L, Qiu X. Utilizing bert for aspect-based sentiment analysis via constructing auxiliary sentence[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics. 2019.
- [69] He R, Lee W S, Ng H T, et al. Exploiting document knowledge for aspect-level sentiment classification[C]//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 2018: 579-585.
- [70] Ziser Y, Reichart R. Pivot based language modeling for improved neural domain adaptation[C]//Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers). 2018: 1241-1251.
- [71] Li Z, Wei Y, Zhang Y, et al. Hierarchical attention transfer network for cross-domain sentiment classification[C]//Thirty-Second AAAI Conference on Artificial Intelligence. 2018.
- [72] Li Z, Zhang Y, Wei Y, et al. End-to-end adversarial memory network for cross-domain sentiment classification.[C]//IJCAI. 2017: 2237-2243.
- [73] Du C, Sun H, Wang J, et al. Adversarial and domain-aware bert for cross-domain sentiment analysis [C]//ACL. 2020.
- [74] Qu X, Zou Z, Cheng Y, et al. Adversarial category alignment network for cross-domain sentiment classification[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). 2019: 2496-2508.
- [75] Zhang Z, Wu Y, Zhao H, et al. Semantics-aware bert for language understanding[C]//AAAI. 2020.
- [76] Zhou J, Tian J, Wang R, et al. Sentix: A sentiment-aware pre-trained model for cross-domain sentiment analysis[C]//Proceedings of the 28th international conference on computational linguistics. 2020: 568-579.
- [77] Levine Y, Lenz B, Dagan O, et al. Sensebert: Driving some sense into bert[J]. arXiv preprint arXiv:1908.05646, 2019.
- [78] Ke P, Ji H, Liu S, et al. Sentilr: Linguistic knowledge enhanced language representation for sentiment analysis[J]. arXiv preprint arXiv:1911.02493, 2019.

- [79] Ma Y, Peng H, Cambria E. Targeted aspect-based sentiment analysis via embedding commonsense knowledge into an attentive lstm[C]//Proceedings of AAAI. 2018: 5876-5883.
- [80] Zhou J, Huang J X, Hu Q V, et al. Sk-gcn: Modeling syntax and knowledge via graph convolutional network for aspect-level sentiment classification[J]. Knowledge-Based Systems, 2020, 205:106292.
- [81] Nguyen T H, Shirai K. Phrasernn: Phrase recursive neural network for aspect-based sentiment analysis[C]//Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. 2015: 2509-2514.
- [82] He R, Lee W S, Ng H T, et al. Effective attention modeling for aspect-level sentiment classification [C]//Proceedings of the 27th International Conference on Computational Linguistics. 2018: 1121-1131.
- [83] Gu S, Zhang L, Hou Y, et al. A position-aware bidirectional attention network for aspect-level sentiment analysis[C]//Proceedings of the 27th International Conference on Computational Linguistics. 2018: 774-784.
- [84] Li X, Bing L, Lam W, et al. Transformation networks for target-oriented sentiment classification[C]//ACL. 2018: 946-956.
- [85] Kiritchenko S, Zhu X, Cherry C, et al. Nrc-canada-2014: Detecting aspects and sentiment in customer reviews[C]//Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014). 2014: 437-442.
- [86] Tang D, Qin B, Feng X, et al. Effective lstms for target-dependent sentiment classification[C]//Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. 2016: 3298-3307.
- [87] Tang D, Qin B, Liu T. Aspect level sentiment classification with deep memory network[C]//Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. 2016: 214-224.
- [88] Chen P, Sun Z, Bing L, et al. Recurrent attention network on memory for aspect sentiment analysis [C]//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. 2017: 452-461.
- [89] Fan C, Gao Q, Du J, et al. Convolution-based memory network for aspect-based sentiment analysis [C]//The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval. ACM, 2018: 1161-1164.

- [90] Zhang M, Zhang Y, Vo D T. Gated neural networks for targeted sentiment analysis.[C]//AAAI. 2016: 3087-3093.
- [91] Wang J, Li J, Li S, et al. Aspect sentiment classification with both word-level and clause-level attention networks.[C]//IJCAI. 2018: 4439-4445.
- [92] Ma D, Li S, Zhang X, et al. Interactive attention networks for aspect-level sentiment classification [C]//IJCAI. 2017: 4068-4074.
- [93] Wang Y, Huang M, Zhu X, et al. Attention-based lstm for aspect-level sentiment classification[C]// Proceedings of the 2016 conference on empirical methods in natural language processing. 2016: 606-615.
- [94] Zhou J, Chen Q, Huang J X, et al. Position-aware hierarchical transfer model for aspect-level sentiment classification[J]. Information Sciences, 2020, 513:1-16.

索引

Aspect Term Extraction, ATE, 25
Aspect-level Sentiment Analysis, ABSA, 22

Comparative Sentiment, 2

Document-level Sentiment Analysis, 9

emotion classification, 7
Emotional Sentiment, 3
Explicit Sentiment, 2

Implicit Sentiment, 3

opinion collocation extraction, 7
opinion collocation polarity classification, 7
opinion holder extraction, 8
opinion identification, 6
Opinion Mining, 1
Opinion Summarization, 8
opinion target extraction, 7
opinion word extraction, 7

polarity classification, 6

Rational Sentiment, 3

Sentiment analysis, 1
Sentiment Classification, 6
Sentiment Information Extraction, 7
sentiment strength detection, 7
Spam Review Detection, 9
Stance Detection, 8
subjective classification, 6

主客观分类, 6

句子级情感分析, 16
垃圾评论检测, 9
属性级情感分析, 22
属性词抽取, 25

常规型观点, 2

情感信息抽取, 7
情感分析, 1
情感分类任务, 6
情感强度判别, 7
情绪, 3
情绪分类, 7
感性感, 3
文档级情感分析, 9
显式观点, 2
极性分类, 6
比较型观点, 2

理性情感, 3
篇章级情感分析, 9
细粒度情感分析, 22

观点, 2
观点持有者, 8
观点挖掘, 1
观点摘要, 8
观点识别, 6
评价对象抽取, 7
评价搭配抽取, 7
评价搭配极性判别, 7
评价词抽取, 7
辩论立场检测, 8

隐式观点, 3