

清华大学学位论文 L^AT_EX 模板

使用示例文档 v5.4.3

(申请清华大学工学硕士学位论文)

培 养 单 位: 计 算 机 科 学 与 技 术 系

学 科: 计 算 机 科 学 与 技 术

研 究 生: 梁 锡 豪

指 导 教 师: 徐 明 星 副 教 授

二〇一九年三月

An Introduction to L^AT_EX Thesis Template of Tsinghua University v5.4.3

Thesis Submitted to
Tsinghua University
in partial fulfillment of the requirement
for the professional degree of
Doctor of Engineering

by
Liang Xihao
(**Computer Science and Technology**)

Thesis Supervisor : Professor Xu Mingxing

March, 2019

关于学位论文使用授权的说明

本人完全了解清华大学有关保留、使用学位论文的规定，即：

清华大学拥有在著作权法规定范围内学位论文的使用权，其中包括：(1) 已获学位的研究生必须按学校规定提交学位论文，学校可以采用影印、缩印或其他复制手段保存研究生上交的学位论文；(2) 为教学和科研目的，学校可以将公开的学位论文作为资料在图书馆、资料室等场所供校内师生阅读，或在校园网上供校内师生浏览部分内容。

本人保证遵守上述规定。

(保密的论文在解密后应遵守此规定)

作者签名：_____

导师签名：_____

日 期：_____

日 期：_____

摘 要

自 Web2.0 普及后，人们习惯在互联网的各种平台上透过不同的媒体来分享和传递他们的想法和情感。透过对这些媒体进行情感分析，我们可以得知人们的想法和态度。譬如透过分析产品评论了解用户对新产品是否满意，又或者分析社交平台上的舆论了解网民对新政策是否同意，以此快速响应，其中蕴含着丰富的商业价值和政治价值，其相关技术的研究价值在近年也得到了重视。而文本作为较常见的媒体之一，加之部分微博等社交平台推出了用户生产数据的采集服务，提供了充裕的文本数据资源，面向文本的情感识别研究得以更快的速度发展。然而随着在互联网上的交互方式变得丰富，文本的数据结构也变得多种多样，如微博上的短文本、在线实时聊天的两人或多人对话、讨论区里的帖子回复。在处理不同数据结构时，该如何提取当中的信息并加以运用来预测目标结果，成为了重要的研究课题。

本文主要探讨了两个不同数据结构下的情感识别问题，并提出了一个通用的文本情感识别系统设计方案。论文的主要内容如下：

- 面向微博的反讽识别。反讽作为一种特殊的修辞手法，起着把文本字面意思反转的效果。识别文本中反讽修辞的使用对于正确理解发言者的意图起着关键性的作用。本文研究了在微博中识别反讽的出现以及其对应类型，深入了解实验数据当中各个类型的样本的特性，分析单个分类器对各个类的识别能力，给出一个基于集成学习的反讽识别框架，以此在多个独立分类器的基础上实现一个集成系统来达到更好的识别能力。
- 面向三轮对话的情感识别。为了更好地识别一段文本的情感，引入上下文信息被认为有指导性作用。而在文本对话的场景下，要识别某个人在某次发言所表达的情感，可以考虑引入该次聊天的历史记录。本文研究了在两人轮流发言的三轮对话中，识别最后一轮发言所表达的情感。此处采用与面向微博的反讽识别相同的系统设计以评估设计方案的泛用性，另外我们进一步探索了上下文信息的运用方面，透过把不同的运用方面体现在分类器的不同设计方案上，并根据其性能差别和对神经网络权重的分析，了解上下文信息如何有助于识别正文的情感。

关键词：意图识别；情感分析；反讽识别；深度学习，集成学习

Abstract

An abstract of a dissertation is a summary and extraction of research work and contributions. Included in an abstract should be description of research topic and research objective, brief introduction to methodology and research process, and summarization of conclusion and contributions of the research. An abstract should be characterized by independence and clarity and carry identical information with the dissertation. It should be such that the general idea and major contributions of the dissertation are conveyed without reading the dissertation.

An abstract should be concise and to the point. It is a misunderstanding to make an abstract an outline of the dissertation and words “the first chapter”, “the second chapter” and the like should be avoided in the abstract.

Key words are terms used in a dissertation for indexing, reflecting core information of the dissertation. An abstract may contain a maximum of 5 key words, with semi-colons used in between to separate one another.

Key words: Opinion mining; sentiment analysis; irony detection; deep learning; ensemble learning

目 录

第 1 章 引言	1
1.1 研究背景与意义	1
1.2 国内外研究现状	2
1.2.1 情感模型	2
1.2.2 情感识别	3
1.2.3 反讽识别	6
1.2.4 社交媒体上的文本意图识别	7
1.3 存在的问题	8
1.4 本论文的内容安排	9
第 2 章 问题分析和研究框架	11
2.1 本章引论	11
2.2 问题分析	11
2.3 研究框架	13
2.3.1 框架入口	13
2.3.2 基础分类器	13
2.3.3 集成识别系统	14
2.4 本章小结	14
第 3 章 意图识别技术	15
3.1 本章引论	15
3.2 文本预处理	15
3.2.1 分词	15
3.2.2 拼写修正	16
3.2.3 规范化	17
3.3 特征提取	17
3.3.1 词嵌入	17
3.3.2 词汇特征	19
3.3.3 句法特征	19
3.3.4 语义特征	19
3.4 模型	19
3.4.1 传统机器学习	19

3.4.2 人工神经网络	21
3.4.3 集成学习	26
3.5 本章小结	27
第 4 章 面向微博的反讽识别	28
4.1 本章引论	28
4.2 本章小结	28
第 5 章 面向三轮对话的情感识别	29
5.1 本章引论	29
5.2 本章小结	29
插图索引	30
表格索引	31
公式索引	32
参考文献	33
致 谢	38
声 明	39
个人简历、在学期间发表的学术论文与研究成果	40

主要符号对照表

BOW	词袋模型 (Bag of Words)
BP	反向传播算法 (Back-propagation)
BRNN	双向递归神经网络 (Bidirectional Recurrent Neural Network)
CRF	条件随机场 (Conditional Random Field)
GRU	门控循环神经元 (Gated Recurrent Unit)
LSTM	长短时记忆网络 (Long Short-Term Memory)
POS	词性 (Part-of-speech)
RNN	递归神经网络 (Recurrent Neural Network)
SVD	奇异值分解 (Singular Value Decomposition)
SVM	高性能计算 (High Performance Computing)
TF-IDF	词频-逆文档频度 (Term Frequency - Inverse Document Frequency)

第1章 引言

1.1 研究背景与意义

意图识别旨在了解人们对特定事件或实体的想法、态度和情感。自 Web2.0 普及后,大量网民每天在互联网生产着不同媒体类型的数据,其中包含各种信息,如个人生活经历,购买行为,对产品服务的体验评价,对社会时事的看法等等。从人们的日常社交需求来看,这种借由互联网媒体的分享非常便捷,我们可以了解到亲朋好友的近况,也或者随时和不认识的网民交流对具体事件的看法。在商业上,借由对用户的网络行为进行分析,企业可以对应客户或者潜在客户有更深入的了解,对他们的需求和反馈及时作出反应将带来战略性的优势。在社会管理上,政府可以透过对网民在网络上的发言了解人民的想法和舆论的走向,进而作出相应的措施。正因为互联网的普及,才使得以上基于对特定人群的了解来进行决策的做法成为可能。随着数据资源变得丰富,相应的技术在近年有明显的发展,目前市面上已经有公司(如国内的腾讯和阿里巴巴,以及国外的微软和亚马逊等)提供基于大型社交媒体平台(如微博,讨论区等)上的数据进行意图分析相关的规范化服务,然而相应的技术依然有进步空间,研究工作还在不同方向上摸索。

情感在人们的思想表达和交流中起着重要作用^[1],比起了解该想法的细节内容,情感对应该想法的大体倾向。譬如在分析用户对新产品的评论时,从对正负性情感反馈的统计可以得知新产品是否能让大部分的客户满意,或者筛选出表示不满意的用户再进行深入分析,因此情感识别是意图分析中属于非常重要的一环。而由于在互联网上,大部分情况下用户以文本表达想法,面向文本的意图识别成为了近年的最要研究课题之一。相对于人们面对面交流的场景,聆听者可以根据发言者的肢体语言,面部表情以及声调变化等额外提示更好地理解发言者所表达的内容,然而这些信息并不存在于文本当中,这也正是对文本进行情感识别本身的难点之一^[2]。

反讽的修辞手法在意图识别当中起着特殊的影响。Henry Watson Fowler 在《The King's English》一书中描述“即使对反讽的定义有数百种,其中只有包含‘表面意思和实际意思不同’这个概念的才能被接受”。Eric Partridge 在《Usage and Abusage》一书中指出“反讽存在于所表达意思的另一面”。总的多说当反讽在文本中出现,那么文本所表达的意思应该和字面表达的意思完全相反。譬如某人表示“我就喜欢你不断挑战我的底线”,从字面上理解应倾向于正面情感,然而根据常识可知“挑战底线”是一种让人反感的行为,与“喜欢”相矛盾。这段文字实则暗示发言者的“底线”

正被“你”“挑战”，表达的是负面情感。因此在意图识别和情感分析当中，正确检查出反讽的使用能够避免对内容的错误理解。

1.2 国内外研究现状

1.2.1 情感模型

情感计算的基础是对情感作出描述，现有的描述方式可以分成两个大类：范畴观和维度观。范畴观即把不同情感对应到一组离散的情感标签上 Alm 等人^[3]，其中具代表性的有 Plutchik 的情感模型^[4]和 Ekan 的情感理论^[5]和。Plutchik 的情感模型包含十种情感：愤怒、恐惧、悲伤、厌恶、期待、信任、高兴、惊讶；这些情感都各自对应具有重要生存意义的行为，各种复杂的情感都是由这此基本情感构成，另外这十种情感可以分成五组对立的情感对。Ekan 在十年后提出的情感理论和 Plutchik 的相似，但相对地少了期待和信任，是一个六类情感模型。

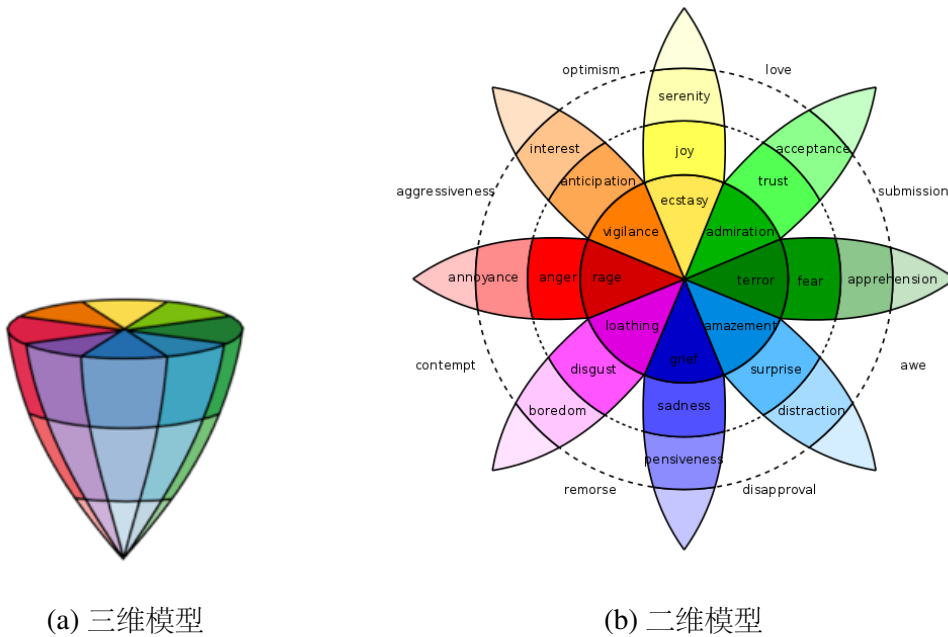


图 1.1 Plutchik^[4] 提出的情感轮模型

以维度观描述情感就是把情感映射到多维空间的点上，而目前维度观情感模型以二维和三维空间的为主。二维情感模型中较有代表性的是 Russell 提出的环状模型 (Circumplex Model)^[6]，其中纵坐标对应情感的激活度 (Arousal)，横坐标对应情感向性 (Valence)，而不同的情感则分布在一个环状的区域内。Bradley 等人^[7]提出的向量模型 (Vector Model) 对横轴和纵轴的定义相似，但其理论假设高激活度的情感应该有较明显的正向或负向，相对地低激活度的情感则偏向中性，故

情感分布在一个回力标形状的区域。Watson 和 Tellegen^[8] 提出的 PANA (positive activation-negative activation) 模型和前两者在理论上则有明显的不同，他们认为情感的正面作用和负面作用是两个独立的成分，所以在模型中纵轴和横轴分别表示情感正面作用和负面作用的强弱，其效果相当于把 Russell 等人提出的环状模型的向量空间旋转 45 度^[9]。

基于三维空间的情感模型中具有代表性的有 Plutchik^[4] 提出的情感轮模型，Plutchik 认为情绪之间包含强度，相似性和两极性三种维度，椎体的顶部和底部分别对应强的情绪和弱的情绪，相似的情感对应椎体中相近的位置，对立的情绪则会对应到椎体中对立的位置上。另外还有 Mehrabian^[10] 提出的 PAD 模型，其三维空间的三个坐标轴分别对应情感愉悦度 (Pleasure)、激活度 (Arousal) 以及优势度 (Dominance)。较近期被提出的是 Hugo^[11] 的情感立方体模型，其三维空间的三个坐标轴分别对应 5-羟色胺 (5-hydroxytryptamine, 5-HT)，多巴胺 (dopamine, DA) 和去甲肾上腺素 (noradrenaline, NE) 三种神经递质所产生信号的强弱，并对空间中一个立方体的八个顶点标记了其对应的情感。

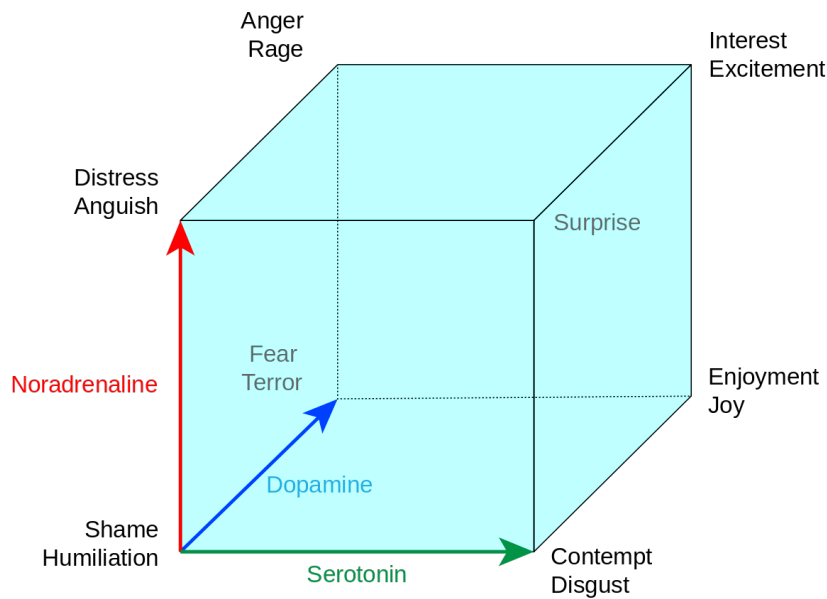


图 1.2 Hugo^[11] 的情感立方体模型

1.2.2 情感识别

对应上述情感模型的分类，情感识别研究可以分成两类。第一类对应范畴观，给定一组情感类型，判断一段文本或针对文本内的某个方面所表达的情感倾向于该组情感中的哪一个，或者是否包含这一组情感中的一个或多个情感。如国际比赛 SemEval2018^[12] 任务一的子任务要求识别一段微博中是否包含愤怒、恐惧、悲

伤等十一种情感中的一种或多种情感。另一类情感识别研究对应维度观,对于给定的情感属性,判断一段文本中该情感属性的强度。其中常见的有情感向性的二分类问题(正性或负性)、三分类问题(正性、中性或负性)、五分类问题(非常正性,正性、中性、负性或非常负性)。其中五分类问题的研究对象一般是互联网上五星评分制的产品评论或者电影评论等。而对目前文本情感识别的研究按照文本的粒度可以大致分成三类:文章级别,句子级别,属性级别。

在文章级别的情感识别中,虽然一篇文章由多个句子组成,均假设其整体存在某种情感偏向,而研究目标则是自动识别出该种情感偏向的类型或强度。**Turney**^[13]利用线上电影评论中"推荐"(大拇指朝上)和"不推荐"(大拇指朝下)的标记研究对电影评论的正负性情感识别。他提出利用点互信息(Pointwise Mutual Information, PMI)来评估单词的语义倾向性,其中PMI透过在大型语料库中统计两个单词的共同出现的情况来评估两个单词的相似度。首先利用PMI值评估各个单词与正向情感的代表性单词"excellent"和负向情感的代表性单词"poor"的相似性,再取这两个PMI值作差得出该单词在语义上倾向于哪一种情感。最近透过计算整段评论的平均语义倾向性评估整体的情感倾向。**Pang**和**Lee**^[14]采用了不同的方法研究相同的问题。他们首先对评论中每个句子的主观程度进行评分,利用评分构造一个带权重的句子关系图,再基于最小割算法结合上下文加强对每个句子主观程度的判断,过滤评论中不带主观情感的句子后再判断整个评论的情感向性,以此加增强了识别能力。**Tang**等人^[15]则研究了产品评论的五级评分预测。除了评论的文本内容,他们进一步引入了用户的信息和产品的信息。经过数据分析,他们发现相同用户对不同产品(此处对应电影)的评论和评分较不同用户之间的一致,另外不同用户对同一产品的评论和评分较不同产品之间的一致,这显示了用户和产品各自都存在一些相对固定的属性。因此**Tang**等人提出一种对用户和产品提取语义特征的方法,并应用于评论的五级评分预测。实验结果显示他们的方法在多个数据集上达到了较好的性能,这同时引出了加入背景信息来加强系统识别能力的可能性。

然而一篇文章有可能同时表达了多种观点和情感,因此有另一类研究针对句子或短文本所表达的情感,即句子级别的情感识别。由于句子的文本长度较文章的短,文本内部的逻辑较简单,但相对地所包含的提示信息也较少,课题的难点与前者有所不同。**Khan**等人^[16]研究了对线上评论中句子进行正负性情感识别。他们首先区分出评论中各个句子的主观性和客观性,然后针对带主观情感的句子,利用开源自然语言工具 SentiWordNet 获取各个单词的正负情感属性,再根据句子的词性标注和他们设计的规则计算整个句子的情感向性。他们的研究默认单词的正负情感属性在不同场景下不变,然而**Li**等人^[17]指出部分单词在特定场景下会有

不同的情感向性，因此他们提出了一种有监督学习方法，自动学习各单词在指定领域下的情感向性，以此作为文本的特征，并应用于对产品评论的情感识别中。实验结果显示使用 SentiWordNet 提供的全局情感评分和针对领域评估的情感评分相比，后者的识别能力更好，同时验证了他们的假设。一些研究则选择端到端地学习单词的情感和语义倾向，在 Santos 和 Gatti^[18] 对电影评论和微博的情感识别研究中，他们提出了利用卷积神经网络分别从字符级别和词级别计算出单词对应的特征向量，然后同样以卷积神经网络结合句子中各单词的向量得出句子的表示向量和预测整体的情感倾向。结果显示他们的方法较早期的其他方法性能更好。

在一些应用场景当中，我们希望了解发言者对某个特定对象或者它的某个特定属性的想法。譬如新手机推出市场后，厂商需要了解用户对手机的续航能力，拍照质量，交互体验等各方面的评价，对于评论中同时谈论手机的多个方面并且好评和差评不一时，应该针对各个属性分别识别发言者所表达的感情，因此有了属性级别的情感识别。Che 等人^[19] 提出一种句子压缩算法，透过对句子进行依存句法分析，过滤与目标属性的情感无关的内容。他们采用了多种语义和语法特征作为输入，以条件随机场 (Conditional random field, CRF) 作为分类器。实验结果显示过滤掉不相关的文本部分后，识别性能有所提升。Wang 等人^[20] 研究了对网上评论中特定实体或属性的情感识别。他们提出了一个基于注意力机制的长短时记忆网络 (Long Short-Term Memory, LSTM)，特点在于只以词向量作为输入，而不采用其他传统的语义和语法特征，另外利用注意力机制自动识别与目标相关的内容。他们实验基于国际比赛 SemEval-2014 任务四^[21] 中的一个子任务，结果然显示他们的系统性能达到了当时的技术水平。另外透过对注意力单元的输出进行分析，验证了注意力机制能有效识别文本中与目标相关的内容。Tang 等人^[22] 分别前述数据集中电脑和餐厅相关的两组样本进行属性级别的情感识别，但有别于当时主流的以特征提取为主的浅层机器学习方法（如支持向量机）和针对序列的深度学习模型（如循环神经网络），他们提出了一个基于注意力机制的深度记忆网络。另外针对文本中各个单词和目标单词在句子中的距离，作者引入了距离信息提出了注意力单元的多种变形。实验结果显示他们提出的深度记忆网络达到了当时最优的系统（手工特征结合支持向量机）的性能。另外对注意力单元的中间结果进行人工分析，验证了在同一段评论中，引入距离信息的注意力单元有助于区分不同单词对不同目标的情感的贡献。

1.2.3 反讽识别

反讽识别技术作为很多自然语言系统的一部分，其应用场景有线上评论分析，人机对话等，对正确理解评论的情感倾向和话语的意图起着辅助的作用。Tsur 等人^{[23][24]}研究了对微博平台 Twitter 上的微博以及电商平台亚马逊上的评论进行反讽强度的识别，按明显不含反讽和明显表示反讽分成五级，由人工进行标注。他们提出的 SASI 算法分别从文本提取了词频相关的模式特征以及基于标点符号的特征，以 K 最近邻算法作为分类器。另外利用在 Twitter 按井号标签 #sarcastic 自动爬取了额外的反讽语料用于初步的模型训练。对实验结果的比较证明了各种特征的有效性以及添加额外语料对模型训练的帮助。

为了以较低成本获取大量的反讽语料，很多研究从微博平台根据井号标签自动筛选出可能带反讽的微博。Reyes 等人^[25]利用 #irony, #education, #humor, #politics 在 Twitter 上自动获取四组英语微博，并把 #irony 对应的微博和另外三组微博两两组成二分类实验。他们提出了四个方面的文本特征以及对应的提取方法，包括：特殊标记（词汇和标点符号等）、不可预期性、表达风格、情感特性；并比较了每项特征在各组微博中的出现情况，显示了与反讽类微博的相关性。另外分别采用朴素贝叶斯和决策树作为分类器，但没有明显的性能区别。

类似的数据收集方法对其他语言同样适用。Kunneman 等人^[26]则利用对应的德语井号标签来获取反讽语料，以此研究德语微博中的反讽识别。他们取 N 元语法作为输入特征，以 Balanced Winnow^[27]作为分类器，在测试集上达到约 0.85 的召回率和 0.87 的 AUC 值。但他们进一步经过人工检验评估了基于井号标签自动标注的有效性，分析结果显示该方法获取的反讽样本中包含约 10

有别于早期以手动设计的特征作为输入和以非神经网络的机器学习方法建模，Poria 等人^[28]首次将神经网络应用于对微博的反讽识别。他们的算法框架主要包含四个卷积神经网络，并分别利用不同的数据集进行预训练，分别对应反讽识别、情感极性识别、情感类型识别和性格识别。最后取四个卷积神经网络的中间结果作为特征，利用支持向量机 (Support Vector Machine, SVM) 进行后融后得出最终预测结果。实验显示引入反讽识别以外的三个语料库提升了系统的识别能力。

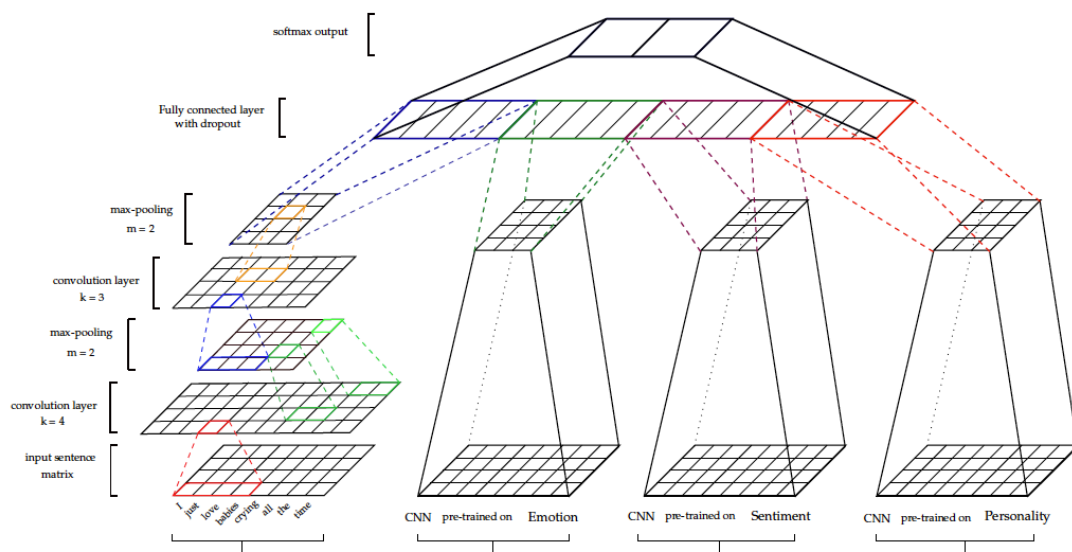


图 1.3 引入多个领域信息的反讽识别神经网络模型，引自 [28]

1.2.4 社交媒体上的文本意图识别

随着社交媒体的普及，网民们习惯于在微博，讨论区等平台上表达个人意见和互相讨论。有别于新闻或学术材料等较正式的文件，网民可以较随心所欲地发言，社交媒体因此成为了意图识别的重要研究对象之一。但和一些产品评论或文章等文本类型不同，社交媒体上的文本普遍较短^[29]，缺少对背景信息的提示，难以判断其内容所属的领域，这对正确理解其表达的想法造成困难。另外语言的使用有别于正式的用法，在中文微博中会出现新的短语或对旧短语有新的解释^[30]，如“锦鲤”暗示“好运”，“灌水”表示“发表没有意义的内容”，在英文微博中会出现错拼字、非正式缩略语、表情符^{[31][32]}，如“tnx”对应英文单词“thanks”，“:)”表示微笑等。虽然没有正式的语言组织对这些新的用法进行整合，但因为这些新用法更方便或对思想的表达更丰富到位，随着在网络上的传播而在网民之间达成了共识。传统的文本特征提取建立在单词标准用法和正规的语法分析基础上，而由于新词汇和新用法的出现，会出现词汇的意思无法被识别或被错误理解，这对于透过人工智能来理解文本成为了一大难点。因此有别于传统的文本研究，在面向社交媒体的文本意图识别时，需要采用额外手段对文本进行预处理进行规范化，或透过大量语料尝试自动学习其语言属性。

Khan 等人^[33]研究了微博的正负中性情感识别。他们提出了一个混合三个分类器的情感识别框架来解决数据稀疏的问题，以此提高整体的分类性能。另外还提出了一组针对微博文本的预处理步骤，其中包括俚语和缩略语分析，词干提取，拼写检查和修正，用户名和井号标签移除等。他们的系统在 6 个微博数据集上达

到了平均 83.3% 的 F1 值以及平均 85.7% 的准确率，和同类型技术比较后验证了他们系统以及预处理手段的有效性。

Angiani 等人^[34] 针对英语微博比较了其种常用的文本预处理技术对情感分析的影响，其中包括单词的规范化、表情符到情感标签的映射、俚语映射、词干提取、停词过滤等。分析显示除了俚语映射以外，其他技术均对情感识别有正面影响，其中部分技术有助于统一拼写相似的单汇，以此关联相同概念的词组。但同时采用所以预处理技术并不保证达到最好的效果，作者指出依然需要根据应用场景和文本的特性做选择。

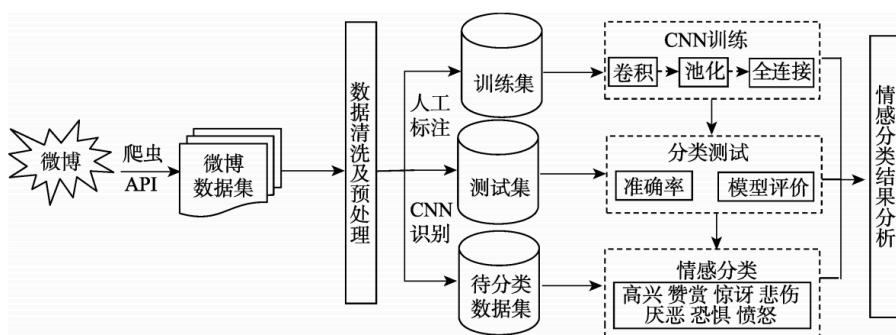


图 1.4 基于卷积神经网络的微博情感分类模型，引自^[35]

张海涛等人^[35] 则研究了中文微博和评论中的文本情感分类。他们针对当时微博上具有一定争议特定话题 # 打呼噜被室友群殴 # 收集数据，确保了样本围绕同一个主题并且有充足的数据量。使用开源工具 NLPIR/ICTCLAS2016 对语料进行分词，再以词向量学习算法 word2vec 从语料中学习词的表示向量作为输入，以基于卷积神经网络的模型作为分类器，另外以支持向量机作为对照算法。实验结果显示在面向原创微博等略长文本的情感识别时，他们的系统比支持向量机性能更好，而面向短文本时则相反。

1.3 存在的问题

由于近年来各企业或机构对意图识别的需求增加，相关技术的研究备受关注，加上深度学习的快速发展，情感识别和反讽识别的性能水平也在逐年提升。然而相关研究依然存在一些空白部分值得进一步探索：

- **对不同数据结构的处理。** 受限于互联网上文本的呈现方式以及数据的收集方式，早期的大部分面向文本的意图识别以文章、微博、评论等形式的数据作为研究对象。随着社交平台功能加强，除了研究的对象本身，我们能获取更多相关的背景信息。以微博平台为例，除了微博的文本内容以外，我们能

根据其历史上发布过的内容分析用户的性格和语言使用习惯，另外可以根据微博下的评论加强对微博本文的了解。但相对地，数据结构变得复杂，该如何从中提取特征，该如何对不同信息之间的相互关系建模，这些问题都有值得尝试的空间。

- **针对优化目标的系统调整。**在不同的使用场景下，针对不同的应用目标，我们会采用不同指标来评估系统的识别水平，并以提高该指标为目标对系统进行优化。而对一个指标的加强往往会导致另一个指标的下降，在数据分布不均匀的分类问题当中，这一影响尤其明显。譬如在反讽识别当中，由于微博上包含反讽的微博明显较不含反讽的微博少，若以准确率作为优化指标，整个系统会倾向将微博判断为不含反讽，含反讽的微博召回率会较低。若改以召回率作为优化指标，整个系统会倾向将微博判断为包含反讽，但因为不含反讽的微博比例较高，会把大量不含反讽的微博误判为包含反讽。近年的相关研究的重点以特征提取、深度学习建模、迁移学习等方向为主，在给定在相同的训练数据以及相同的模型框架下，模型的拟合能力是有限的，如何针对不同的优化目标对系统进行调整是个极少被探讨的问题。

1.4 本论文的内容安排

本文提出了一个面向文本意图识别的系统设计方案，并将其应用于两个数据结构不同的意图识别问题：面向微博的反讽识别、面向三轮对话的情感识别。论文的工作主要围绕我们提出的系统设计方案和这两个应用场景展开。

在论文的第二章，我们会对面向文本的意图识别一类的问题作出分析，并给出一个统一的数学定义。基于该定义，我们将给出一个通用的意图识别系统框架，细节说明其中每个模块的功能和目的。该框架将被具体应用到后面章节的实验当中。

在第三章，我们会介绍在意图识别系统中使用到的自然语言处理技术和意图识别技术，并对应到前一章描述的系统框架的各个模块当中。我们会对后续实验中使用到的技术给出相对充分的说明，同时也会相关的技术给出概要的描述，以便于其他研究者在本论文未深入探索的方向作出拓展性的研究。

第四章中，我们将把前述的系统框架和技术应用于面向微博的反讽识别，透过对数据的特性进行人工分析，我们将给出具体的反讽识别系统实验方案，透过对细节的调整和性能的比较，验证我们采用的策略的效果并分析其中的原因。整个实验基于国际比赛 SemEval2018 的任务三^[36]，其中包含两个子任务。子任务一为二分类问题，要求识别微博是否包含反讽的修辞手法。另一个子任务为四分类

问题，数据与子任务一相同，但原本"反讽"一类的样本被细分成反讽的三个类型：基于相反语义的言语反讽、其他言语反讽、情景反讽。我们将采用官方提供的训练数据和测试数据，并按照比赛的要求开发识别系统，以此和其他参赛系统的性能进行比较，评估我们系统的整理水平。另外透过对错判的样本进行人工检查来进行错误分析，尝试给出系统对该问题的改进方向。

在第五章中，我们将采用与前一章相同的流程来研究面向三轮对话的情感识别。本章的实验基于国际比赛 **SemEval2019** 的任务三^[2]，比赛要求参赛系统识别两人轮流发言的三轮对话中最后一轮发言者所表达的情感，一共四个情感类别：高兴、悲伤、愤怒，其他。除了如前一章中验证系统框架的有效性以外，本章我们将深入分析复杂文本数据结构的不同建模方法对系统最终识别能力的影响，并透过对网络权重和中间结果的分析了解上下文中哪些信息对最终的情感识别起著作用，以及神经网络是否能捕捉到文本对情感识别起关键作用的成分。

第2章 问题分析和研究框架

2.1 本章引论

随着互联网上不同类型的平台出现,人们每天在各种平台上产生着各种各样的行为和发言。如现在微博平台上会对热门时事设置井号标签,人们透过加上对应标签来表达对该事件的想法,以及透过对这些发言进行赞点、分享或评论来表达支持或反驳。利用井号标签收集数据并进行意图分析可以得知网民对该事件的舆论方向。在其他平台上的各种行为记录同样可能有挖掘其意图的价值,意图识别的应用场景也变得多种多样,但即使数据的内容和结构不同,问题的本质是相似的。因此在本章,我们将首先针对意图识别进行分析,并给出统一的形式化表示。

再进一步,基于该形式化表示,我们提出一个面向社交媒体文本的意图识别研究框架。从原始文本的输入到最终识别目标的输出,理清其中每个步骤的功能和目的,给出一个完整识别系统的设计方案,为解决后续章节中研究的问题准备一个统一的切入点。

2.2 问题分析

本小节中,我们将对意图识别中涉及各个元素作出分析,并给出统一的形式化表示来描述他们的相互关系。

不同意图识别问题中都有要被识别的意图倾向 C 。如 Tang 等人^[15]的情感极性识别研究, C 对应需要五级的情感极性。在刘丹丹等人^[37]的微博情感分类研究中, C 对应喜好、安乐、惊奇、厌恶、悲哀、愤恨、恐惧。邓钊等人^[38]的中文反语识别研究中, C 则对应是否包含反讽。

其次是研究主体 T ,它是行为发起者 S 的行为记录,可以认为他表达想法和情感的载体。另外一些场景下会有背景信息 B ,对应所有有助于正确理解 T 的信息。譬如要研究讨论区上帖子的意图倾向,那么 T 就是帖子的内容,包括其中的文本内容、图片、文件附件等, S 则是帖子的发布者。而帖子所在讨论区的类型有助于定位帖子对应的领域,发布者的发布历史显示发布者的一些态度倾向,帖子下的评论从侧表反映主帖的内容,这些就是背景信息 B ,都可能是理解帖子内容的提示。又以 Zahiri 等人^[39]对电视剧台词的情感识别研究为例,那么 T 指电视剧中的台词, S 则是发出这段台词的对应角色, B 则是台词的上文,其中包含其他角

色正在谈论的内容，这些角色和发言者的关系是什么，说话的氛围如何等等，都能对台词的内容有更明确的定位。

意图识别假设对于任意一个研究样本 $t \in T$ ，在给定背景信息 $b \in B$ （或某些情况下假设与背景信息无关），其承载想法或情感必然存在对应的倾向 $c \in C$ 。意图识别首先要从样本主体和背景信息中分别提取出与识别目标相关的信息 f_t 和 f_b ，即需要提出两个映射函数， F_T 和 F_B ，满足 $f_t = F_T(t)$ 和 $f_b = F_B(b)$ 。再进一步根据相关信息识别出其想法或情感倾向，即找出一个映射关系 F_C ，使得 $c = F_C(f_t, f_b) = F_C(F_T(t), F_B(b))$ 。

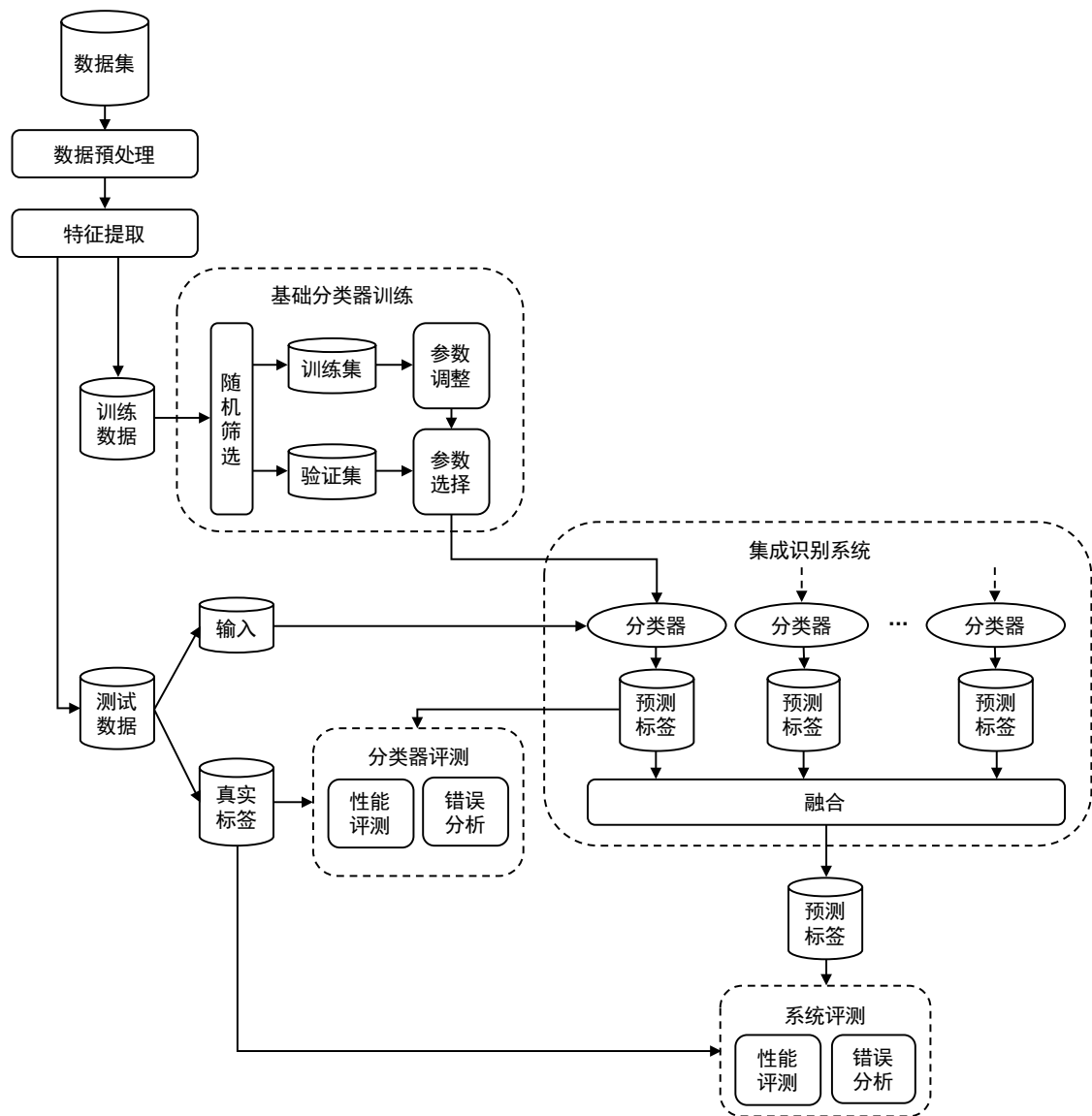


图 2.1 意图识别的研究框架

2.3 研究框架

在本小节，我们将根据前述的形式化表示给出一个意图识别的研究框架。由于在本论文的后续研究均面向英语社交媒体文本的分类问题，我们会针对此研究方法对框架进行细化，从原始文本的输入到最终识别目标的输出，理清其中每个步骤的功能和目的，给出一个集成系统的设计框架。针对不同的应用的场景，在后续实验章节中将给出具体的实现方案。

2.3.1 框架入口

图 2.1显示本文面向意图识别的研究框架。整个框架的输入是实验的数据集，数据集的每个样本包含三个元素，分别是研究的主体、背景信息和意图的倾向。参考前一小节的形式化表示，一个样本可以表示为 $\langle t, b, c \rangle \in T \times B \times C$ 。不论是训练数据或测试数据，他们都采用相同的数据预处理技术和特征提取技术来获取样本的特征，以用作分类器的输入，即透过 F_T 和 F_B 分别获取 $f_t = F_T(t)$ 和 $f_b = F_B(b)$ 。

2.3.2 基础分类器

在特征提取完成后，识别模型的输入就准备好了。接下来首先对基础分类器进行研究和分析，比较不同模型和不同参数在对应问题上的整体性能。对于机器学习方法，在模型的训练过程中，需要将训练数据会分成训练集和验证集两个部分。训练集用于调整模型的参数，根据预先设计的损失函数计算模型当前的预测结果和实际标签的偏差，透过微分计算出参数的调整方向，并以一定的学习步长逐步迭代。由于以上训练过程在数学逻辑上是以达到训练集上的最优解来调整参数，会出现对训练样本过拟合的问题，因此会采用验证集来评估模型对训练集以外的样本的识别性能。在每一轮参数调整后计算当前模型在验证集上的性能指标，最后选择其中最优的一轮参数作为该次模型训练最终的参数。

除了分析不同模型在对应问题上的整体性能，透过对被分类器错误识别的样本进行观察，我们需要分析单个分类器的性能局限在哪，分类器对哪些类别的样本有更高的正确率或召回率，哪些样本之间容易出现混淆，或者在输入数据当中是否存在一些有关键信息的特征没有被成功提取。特别地意图识别的研究课题中，我们会关注数据集标签的正确性。在实验的过程中我们假设标签是正确，但有研究指出自动标注会引入噪声^[27]，人工标注则难以避免地存在主观性。虽然对不同算法之间的性能比较影响不大，但对于技术水平的评估有其分析的价值。

2.3.3 集成识别系统

经过对基础分类器的分析后，我们对不同模型的识别性能有了深入的了解。为了进一步达到更好的性能，下一步是基于基础分类器研究集成识别系统的策略和设计。我们提出了一种集成系统的设计方案，结合多个模型相同的分类器的预测结果来作出最终判断，旨在只使用一种模型的基础上尽可能提升识别的性能，或针对特定的性能指标对整个系统的识别倾向进行调整。参考决策树，其识别过程可以认为是一系列的决策，每一步决策都建立在前一步的决策结果上，并且只关注原始问题中的子问题。类似地，我们提出的集成系统经过多次决策来对样本进行意图识别，第一步先由一个分类器或多个分类器的预测结果融合得到一个基础的预测结果，接下来的每一步需要先对前一步的混淆矩阵进行分析，根据哪些类别的样本被误判对指标的影响较大，设定一个子识别任务，从训练数据中筛选出相关的子集来训练新的基础分类器，得出针对该子识别任务的决策结果，决定是否修改上一步中的预测结果。这一方法的原理在于相同的模型在相同配置下的拟合能力是有限的，而子识别任务是原始问题的简化，模型可以专注于拟合一个局部问题，并对前一步的预测结果提出修改意见，起着补丁的作用。具体方法将在后续实验章节中给出详细说明。

同样地，我们需要分析集成系统的性能。在测试数据上，观察系统经过每一轮决策调整之后的识别性能变化，验证系统框架的设计是否满足假设。另外深入观察混淆矩阵的变化，分析各个子识别任务有效的原因，判断其适用和不适用的场景。

2.4 本章小结

在本章我们首先对意图识别给出统一的形式化表示，引出了其中涉及各个元素并描述他们的相互关系。然后描述了我们面向意图识别的研究框架，其中包括了两个研究重点。一是研究以不同算法得到的分类器对研究问题的整体性能，并作深入的分析，包括单个算法对不同类别样本的识别能力、比较不同算法在不同指标上的区别以理解他们的拟合倾向、观察被错误识别的样本并找出其原因等。二是研究基于基础分类器的集成系统，鉴于单个分类器的识别能力有限，我们提出了一个集成系统的设计方案，旨在结合多个基础分类器的预测结果来得出最终的判断，从而充分发挥一个数学模型的拟合能力，细节实现方法将在后续具体的应用场景中给出说明。

第3章 意图识别技术

3.1 本章引论

根据前一章中描述的研究框架，在本章中我们将针对每一个功能介绍对应的技术实现，其中包括文本预处理、特征提取、分类算法、集成学习。对后续实验中使用到的技术给出相对充分的说明，同时也会相关的技术给出概要的描述，以便于其他研究者在本论文未深入探索的方向作出拓展性的研究。

3.2 文本预处理

文本的预处理是所有面向文本的研究的第一步，其目的是为特征提取做好准备。良好的预处理策略可以在尽可能不掉失重要信息的情况下对样本数据进行简化，增加样本之间重复的模式，减轻模型对数据进行拟合的负担，同时更有效地找出数据之间的相互关系。错误的预处理策略会掉失具有区分能力的信息，甚至产生具有误导性的样本。对于不同的语言，由于其天然的性质不同，预处理的方法也会各异。本小节将针对社交媒体上的英语文本介绍几种预处理技术。

3.2.1 分词

为了从文本提取词级别的特征，我们需要首先将句子切成多个词的序列。

虽然对于英语及大部分欧洲语言，空格隔开的字符必然属于不同的词组，但对于不以空格分隔词组的语言（如中文），分词的作用尤其重要。在某些情况下，分词的结果会影响对句子的理解，如将“乒乓球拍卖完了”切割成“乒乓球拍-卖-完-了”或“乒乓球-拍卖-完-了”，对主体应该是“乒乓球拍”还是“乒乓球”，动词应该是“卖”还是“拍卖”，仅凭字面意思无法确定发言者想表达的意思。特别地社交媒体平台上，新词不断的出现，要正确进行分词就有其独特的难点。而分词并不只针对语言中的单词，还针对标点符号或其他特征字符组成的有特殊语义或情感的字符组合，如现今社交媒体上普遍用多个字符拼接成颜文字，其中最常见的微笑的表情“:)”和伤心的表情“:(”，但如果在分词过程中把前者分割成“:”和“)”就会失去其所带的正向情感信息，这在短文本的情感识别中非常关键。

虽然利用空格和标点符号在大部分情况下可以完成对英语句子的分词，但在一些情况下，标点符号作用为词组的一部分而不是词组的分隔符，而在社交媒体上会有空格被省略的情况，以下是一些需要额外处理的情况^{[40][41]}。一是带句号

的缩写，如“U.S.”、“.com”，句号应该作为词的一部分，或按句号切割“U.S.”将失去其语义。二是具有一定格式的带标点符号的词组，如电子邮箱地址（如 `example@email.com`）、时间（如 `Jan 6th`、`06/01/19`）、电话号码（如 `(123)456-7890`）、网页地址（如 `www.example.com`）等，这些内容在大部分情况下可以被视为一个个体，我们更优先关注这个个体对应什么类型的事物，而不是其细节，譬如将一个句子中的电子邮箱地址或电话号码作替换并不会改变其情感的表达，但识别出一个字符串对应的事情类型，并在清楚它对意图识别没有关联的情况下对其忽略是有意义的。第三种情况是附属词，如“`t`”对应“`not`”，只有正确识别“`’`”的作用才能识别出否定的意思，否则句子的意思将完全相反。值得注意的是，对社交媒体平台上的文本，除了以上在正规英语中会出现的情况，还有出现其他特殊情况，如“`Y!E!S!`”和“`N!O!`”。这需要对数据首先进行人工观察找出特殊的模式，再对语料库进行统计判断其出现频率，若出现频率较高则新增处理规则将对应模式做转换（如将“`Y!E!S!`”替换成“`YES!!!`”）或对问题无关的模式忽略（如国外的微博以“`RE:`”开头表示回复，并不包含任何情感）。

3.2.2 拼写修正

在处理较正式的文件时，我们一般可以默认其文本满足正规的语言语法，词汇基本都是正确的拼写或故意设计的新词（如杂志专栏作家首创针对英国退出欧盟首创单词“`Brexit`”）。不过在社交媒体上，拼写不符合传统英语的情况则非常普遍，这些情况可以分成三大类。

第一类是非刻意的拼写错误，由于社交媒体上的文本普遍是非正式的，用户不会刻意去保证文本的语言正确性，他们更关注于表达自己的想法、感情或其他意图。拼写修正技术一般基于统计的方法，对于一个不存在的词汇，尝试对其拼写进行有限次编辑，再评估编辑后的词在其上下文中出现的概率，最后选出可能性最高的一个词作替换。拼写修正技术已相当成熟地被应用于搜索引擎和手机键盘等应用中，读者可参考相关研究^{[42][43]}了解其细节。

第二种情况是网络上常用的代替用词，如以“`thx`”、“`tnx`”代替“`thanks`”，以“`k`”代替“`ok`”，以“`cant`”代替“`can't`”。虽然这些用法没有被认可为标准用法，但由于方便而在网络上被传播开成功网民们都能理解的用词。对此应对方法有两种，一是人工建立映射表，把代替用词映射到标准的英语词汇，优点在于可以直接引用标准词汇的语义信息，但缺点是需要人工参与，特别是在网络上新用词不断出现的情况下需要持续的更新。另一种是不做处理，利用机器学习方法从大型语料库中自动学习出它的语义，优点在于省去了人工的部分，缺点在于单词的出现很稀疏，对

语料库的依赖很强。

第三种情况是语气加强，如“yeeeees”、“AMAZING”，但对于这一类情况处理的重点不只在于转换成标准用词，而是根据具体的应用问题判断是否要保留这个词存在语气加强的这个信息，譬如在情感识别中，语气加强一般提示了此处的情感需要注意。Baziotis 等人^[44]的做法是在单词前后添加标签示意，如“yeeeees”替换成“yes <enlongated>”，“AMAZING”替换成“amazing <allcaps>”。

3.2.3 规范化

在章节3.2.1中曾提及在一些特定的应用场景下，电子邮箱地址和日期等内容的细节其实对我们的研究任务并不起作用。对应地我们可以将内容一致但拼写不同的字符串转换成同一个词组来表示相同的概念，如 Kouloumpis 等人^[45]在处理微博文本时将网页地址和“@ 用户名”分别用同一个统一的字符串代替。更进一步地，Baziotis 等人^[44]在研究情感识别时，将表达同一种情感的不同表情符替换成对应的情感标签，如“:)))”替换成“<happy>”，“:-D”替换成“<laugh>”。但需要注意替换的内容是否会导致信息掉失，如 Joshi 等人^[46]在针对反讽识别的研究中指出一些反讽透过具体数字的对比来表达，在这种情况下将具体数字统一替换成表示数字的标签可能会导致预处理后的样本失去原本的反讽属性。

3.3 特征提取

数据经过预处理后，我们需要从获得的字符串或字符串序列中提取任务相关的特征，并进行量化方可作为分类器的数学模型的输入。本小节将介绍几种主流的文本特征提取技术。

3.3.1 词嵌入

词嵌入即捕足“词”的语义、语法、情感等信息并将其以向量表示。其中“词”对应预处理中分词得到的词，或对于字符级别的模型，“词”可以对应单个字符^[47]。主流的词嵌入算法分成两大类：基于人工神经网络的方法和基于矩阵分解的方法。

基于人工神经网络的方法假设句子中词之间存在某种上下文关系，并以神经网络建模学习，其中最具代表性的是 CBOW 和 Skip-gram 两种算法。CBOW 的建模原理是句子中的一个“词”可以由它在句子中的前后有限个“词”决定，Skip-gram 的建模原理则是句子中的“词”可以推测出它在句子中的前后有限个“词”。这两种算法都在自然语言处理中常用的词嵌入学习工具 Word2vec^[48]中提供，又由于训练优质的词嵌入模型（指一组词到词向量的映射）需要收集大量的数据和耗费相当的

时间，一些利用 Word2vec 训练好的词嵌入模型会作为公开的资源供研究员使用，以便于其他自然语言相关的研究，更方便于相互复现研究结果。其中较常用的是谷歌提供的模型^①，利用谷歌新闻作为语料训练出复盖约三百万个词的三百维向量集合。然而词向量所包含的信息和语料库紧密相关，所有也有一些研究会按照自己课题的特性自行收集相应的语料来训练词向量。如 Baziotis 等人^[44]利用微博平台 Twitter 上收集的消息训练词向量模型并应用于面向微博的反讽识别，其词向量模型也被公开供其他研究者使用^②。

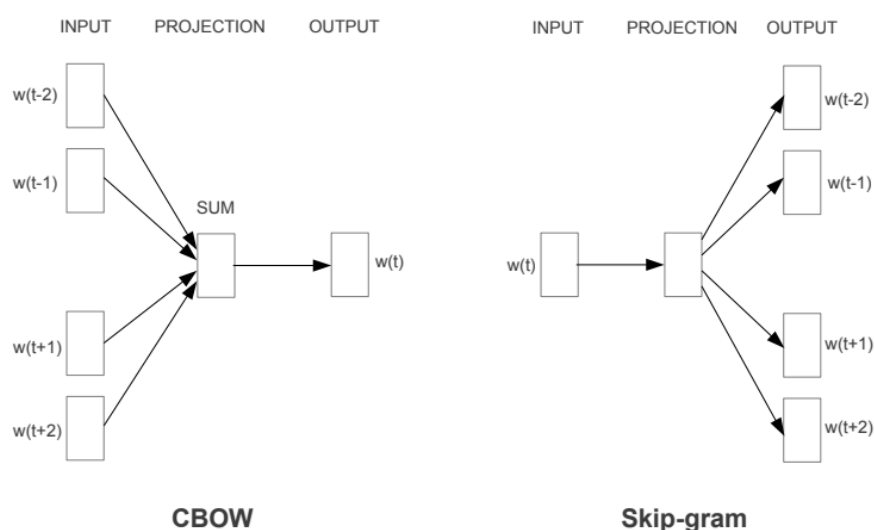


图 3.1 Word2vec 提供的两个算法的模型，引自^[48]

基于矩阵分解的方法假设一个表示词和词或词和文档之间关系的矩阵可以表示为由词向量组成的矩阵和其他矩阵的相乘，因此对该关系矩阵以对应方式进行矩阵分解即可以得到词向量。具代表性的方法是隐藏语义分析方法 (Latent semantic analysis, LSA)^[49]，以词在各个文档中出现的频率得出词-文档矩阵，并假设其文档可以表示为词-词概念、词概念-文档概念、文档概念-文档三个矩阵相乘，利用奇异值分解 (Singular Value Decomposition, SVD) 得到词-词概念矩阵，取每一行作为对应词的词向量。另外还有 Pennington 等人^[50]提出的 GloVe 算法，该算法根据不同词在语料中出现在各自上下文窗口中的频率得出词和词之间的关系矩阵，再分解成两个词-词概念矩阵的相乘。其研究结果显示在部分自然语言处理任务中，采用 GloVe 较 Word2vec 及另外几种词嵌入技术得出的词向量能达到更好的效果。他们同样公开了多个利用不同语料库训练好的词向量模型供其他研究者使用^③。

① <https://code.google.com/archive/p/word2vec/>

② <https://github.com/cbaziotis/ntua-slp-semeval2018>

③ <https://nlp.stanford.edu/projects/glove/>

3.3.2 词汇特征

词汇特征表示文本中一元或多元语法的分布情况，不考虑词的具体意思。譬如常用的词袋模型（Bag of words, BOW），其将一段文本映射到一个固定长度的向量，向量上每一维对应一个一元或多元语法的某种属性，譬如它是否在文本中出现（即独热编码），或它在文本中出现的频率，或其词频-逆文档频度（Term Frequency - Inverse Document Frequency, TF-IDF）。或更简地以单词数量或字母数量作为一维的特征。

3.3.3 句法特征

句法特征表示文本中各个词在句子的语法作用。常见做法是先对文本进行词性（Part-of-speech, POS）标注和依赖树分析，再基于标注提取特征，如形容词的数量、副词的数量，或以标注代替单词后采用词袋模型得出固定长度的特征向量。一些研究会针对性地手工设计句法特征，这要求研究者对语法的了解和挖掘它对研究的问题是否存在关联性，但特征的作用更直观，解释性更强。

3.3.4 语义特征

语义特征基于文本的字面意思，表示文本表达的内容或内容的个别属性（如主题分布），这更接近于人们透过理解句子内容来进行识别。常见技术譬如基于文本中各个词所对应词向量计算句子的表示向量，或利用词向量计算句子中各单词和情感词的相似度来间接评估词的情感极性，或利用聚类算法对语料库的单词进行聚类来获得单词的隐藏语义类别，再计算句子中各类别单词的分布。

3.4 模型

本小节将给出搭建意图识别系统相关的模型。目前用于意图识别的主流模型可以分成两大类：传统机器学习和人工神经网络。我们会分别对这两类模型中具代表性的例子给出介绍，然后再介绍集成学习方法，以在有限个模型的基础上搭建集成识别系统，达到比单个模型更高的性能。

3.4.1 传统机器学习

支持向量机（Support Vector Machine, SVM）^[51] 由 Cortes 和 Vapnik 在 1995 年提出，是目前最常用的传统机器学习算法之一。在对二分类问题的建模过程中，每个训练样本被映射到一个空间中的一个点，而支持向量机尝试把找出一个边缘把空间分成两个子空间（如以直线分割二维空间，以平面分割三维空间），每一个子

空间对应一个类，使得训练样本落在对应类别的子空间中，并且各样本到边缘的距离最大。对于多分类问题，主流做法是把原问题拆解成多个二分类问题。后随着核函数的引入，将支持向量机拓展成非线性的数学模型，大大加强了其建模能力。除了分类问题，支持向量机同样可以应用于回归问题，Drucker 等人^[52]则提出了支持向量机的回归模型。支持向量机最大的限制之一是只接受固定长度的向量作为输入，要求时先从输入数据提取出固定长度的特征向量，换言之基于支持向量机的识别系统非常依赖对特征的提取。

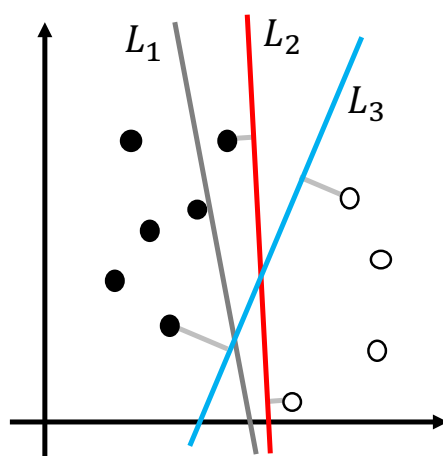


图 3.2 支持向量机在处理二类问题的示意图。其中黑点和白点分别对应两个类别的样本， L_1 不能对两类样本作区分， L_2 和 L_3 均成功区分两个类别的样本，但 L_3 和两个类别样本的距离更大，为更优解

决策树 (Decision Tree) 是另一种常见的机器学习方法。对于分类问题，决策树基于训练数据建立一个树状模型，其中每个非叶子结点对应一个决策，叶子结点对应要识别的类别。在识别阶段，识别过程从根结点开始，经过每个非叶子结点根据输入数据得出决策结果，不同的决策结果会对应下一级的一个结点，如此不断前进直到到达叶子结点，以其对应类别作为分类结果。和支持向量机一样，决策树的输入为固定长度的向量，要求对应的特征提取。现今的研究更多地会使用由多个决策树组成的随机森林 (Random Forest) 而非单个决策树作为分类器，随机森林属于集成学习一类，将在后续章节 3.4.3 中介绍。

3.4.2 人工神经网络

人工神经网络源自仿生学，以数学模型模拟生物神经网络的响应过程和学习过程。**Rosenblatt** 提出的感知器^[53]是最早的人工神经网络，也是现在人工神经网络的基本单元（又称为神经元）。**Rumelhart** 等人^[54]在后来提出了一种多层前馈神经网络，**BP** 网络，由多层的感知器和激励函数反复堆叠而成。更重要的是他们提出了它的学习算法，反向传播算法（**Back-propagation, BP**），当网络的输出结果和预期结果不同，透过构造损失函数计算偏差，再根据损失函数和偏差计算出神经元权重的梯度，其对应了使误差扩大的权重方向，以其反方向作为权重修改的依据。同时将该层输入的修改量作为前一层输出的偏差，逐层重复修改权重，整个过程重复迭代直到误差满足一定条件时学习过程结束。反向传播算法是目前人工神经网络模型训练中的核心算法，但随着网络层数增加，梯度以指数级别增大或缩小，会导致梯度大小过大或近似于零的情况，即梯度爆炸和梯度消失。直到 **Hinton** 等人^[55]提出深度信念网络和无监督逐层训练的策略，其核心思想是在每一轮训练只调整网络中一层的权重，透过固定前一层的输出作为本层的输入，避开了多次传播带来梯度的指数级变化。如此网络逐层经过调整后，再对整个网络进行权重调重，理论上此时的梯度较小，避免了梯度爆炸的出现。这使得现在复杂人工神经网络的学习成为了可能，深度学习正式进入高速发展的阶段。接下来我们将对主流的人工神经网络单元作出介绍。

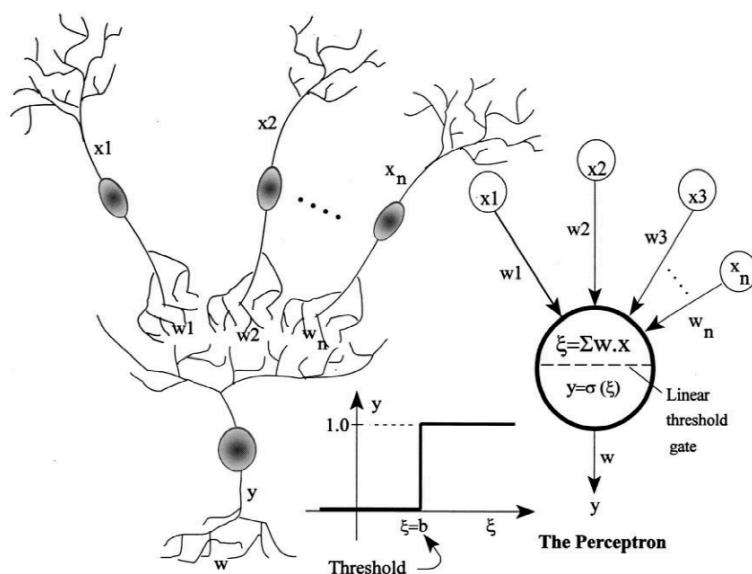


图 3.3 生物神经元与感知器的模型示意图，引自^[56]

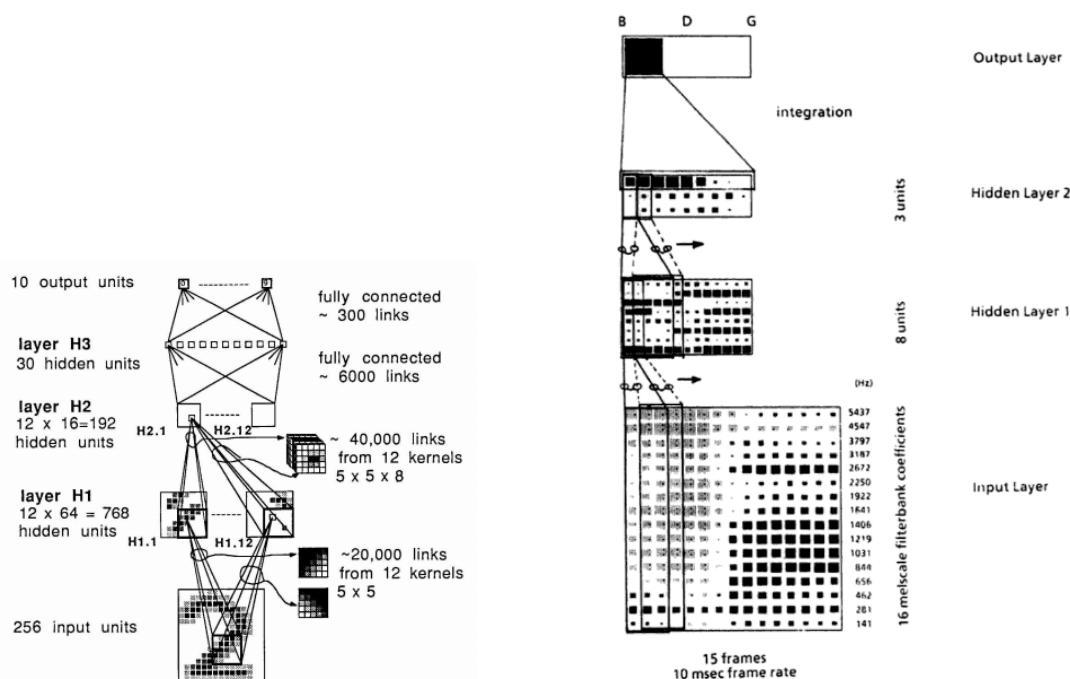
(a) Lecun 等人^[57]的反向传播网络(b) Waibel 等人^[58]的时延神经网络

图 3.4 卷积神经网络模型

3.4.2.1 卷积神经网络

卷积神经网络 (Convolutional Neural Network, CNN) 是目前最常用于计算机视觉的一类人工神经网络, 它的设计来源自生物的视觉皮层。Hubel 和 Wiesel^{[59] [60]} 在对猫和猴子的视觉皮层进行观察时发现其中一些神经元只对接收到的视觉画面的一个区域作出反应, 该区域称为该神经元的感受域 (Receptive Field) 而相邻神经元的感受域会有所重叠, 联合构成对整个视觉画面的接收和反应。这些神经元中再细分为两类, 第一类神经元对画面中特定方向的刺激反应最强烈, 另二类神经元则较前一类神经元的感受域大, 但其反应对刺激在感受域中的具体位置不敏感。

基于以上生物特性, Fukushima^[61] 提出了神经认知机 (Neocognitron), 其数学模型中包含了两个关键结构, 卷积层和降采样层, 与前面两类神经元的功能一一对应。受启发于神经认知机, Lecun 等人^[57] 提出了一个三层的反向传播网络, 包含特征映射和降采样两种功能的元件, 同样与视觉皮层中的两类神经元对应。该网络被应用于手写数字的识别, 和当时同类型网络的性能相比达到了近 30% 的提升, 该数学模型自此成为了现今卷积神经网络的原型。

虽然卷积神经网络起源于视觉处理, 但其变型还应用于其他领域。如 Waibel 等人^[58] 提出的时延神经网络 (Time-Delay Neural Network, TDNN) 比应用于计算机听觉的音素识别, 其核心思想是将每一个时刻的音频特征按时间顺序堆砌, 以

时间上的相邻类比视觉画面上位置的相邻。Zhang 等人^[62]提出了一个字符级别的卷积神经网络并应用于文本主题分类，以字符类别语音中的信号，首次将卷积神经网络用于自然语言领域。

3.4.2.2 递归神经网络

递归神经网络（Recurrent Neural Network, RNN）是另一个人工神经网络的大类，和前馈神经网络最大的区别在于递归神经网络中以某种机制实现了记忆能力来保留前一时刻的状态，因此可以对一个序列输入中前后出现的内容进行关联，换言之对序列的数据进行建模。递归神经网络会在每一时刻接受一个输入，并结合其“记忆”计算这一时刻的输出，故递归神经网络的输出是和输入序列长度相同的序列。递归神经网络的应用场景包括视频内容识别、文本翻译、股票指标预测等。较早期被提出的递归神经网络模型是 Elman^[63] 的 Elman 网络，其数学模型如下：

$$h_t = \sigma_h(W_h x_t + U_h h_{t-1} + b_h) \quad (3-1)$$

$$y_t = \sigma_y(Y_y h_t + b_y) \quad (3-2)$$

其中 $x_t \in \mathbb{R}^d$ 为 t 时刻的输入， $h_t, h_{t-1} \in \mathbb{R}^h$ 分别表示网络在这一时刻和上一时刻的隐藏状态， $y_t \in \mathbb{R}^y$ 为网络的输出， σ_h 和 σ_y 为激活函数， $W_h \in \mathbb{R}^{h \times d}, U_h \in \mathbb{R}^{h \times h}, b_h \in \mathbb{R}^h, Y_y \in \mathbb{R}^{y \times h}, b_y \in \mathbb{R}^y$ 为模型参数。可见在模型中因为引入 h_{t-1} 作为当前时刻的输入而表现出某种记忆，上一时刻的输入 x_{t-1} 间接影响了这一轮的输出。

目前常用的递归神经网络之一是 Hochreiter 和 Schmidhuber^[64] 提出的长短期记忆模型（Long Short-Term Memory, LSTM），其数学模型如下：

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \quad (3-3)$$

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \quad (3-4)$$

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \quad (3-5)$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \sigma_c(W_c x_t + U_c h_{t-1} + b_c) \quad (3-6)$$

$$h_t = o_t \circ \sigma_h(c_t) \quad (3-7)$$

其中 $x_t \in \mathbb{R}^d$ 为 t 时刻的输入， $h_t, h_{t-1} \in \mathbb{R}^h$ 分别是这一时刻和上一时刻 LSTM 的输出（或当 LSTM 作为一个人工神经网络中其中一层时称为隐藏状态）， $c_t \in \mathbb{R}^h$

提当前时刻记忆, $i_t, f_t, o_t \in \mathbb{R}^h$ 分别控制了这一时刻的输入、上一时刻记忆和上一时刻的输出在这一刻的起的作用, 因此称为“输入门”、“遗忘门”、“输出门”的激活向量, $\sigma_g, \sigma_c, \sigma_h$ 为激活函数, 特别地 σ_g 一般采用 *sigmoid* 函数, σ_c, σ_h 一般采用 *tanh* 函数, $W \in \mathbb{R}^{h \times d}, U \in \mathbb{R}^{h \times h}, b \in \mathbb{R}^h$ 为网络参数。和 Elman 网络比较可见, LSTM 除了有 h_{t-1} 传递上一时刻的信息以外多了一个记忆单元 c , 具有更强的记忆功能。在 LSTM 的设计中包含了“输入门”、“遗忘门”、“输出门”和记忆单元, 对序列数据有较强拟合能力, 相对地对于小规模的序列数据容易出现过拟合。

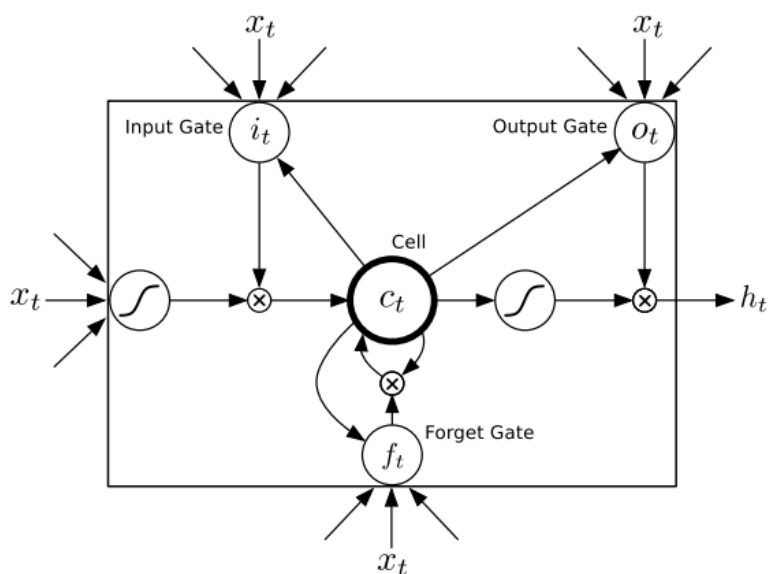


图 3.5 长短时记忆模型, 引自 [65]

为了解决容现出现过拟合的问题, Cho 等人 [66] 提出了门控循环神经元 (Gated Recurrent Unit, GRU)。可以认为它是 LSTM 的一种轻量级变型, 其数学模型如下:

$$z_t = \sigma_g(W_z x_t + U_z h_{t-1} + b_z) \quad (3-8)$$

$$r_t = \sigma_g(W_r x_t + U_r h_{t-1} + g_r) \quad (3-9)$$

$$h_t = (1 - z_t) \circ h_{t-1} + z_t \circ \sigma_h(W_h x_t + U_h(r_t \circ h_{t-1}) + b_h) \quad (3-10)$$

可见在数学模型上 GRU 和 LSTM 非常想似, 虽然 GRU 中省去了专门的记忆单元, 但同样采用了“门”来控制当前时刻的输入和前一时刻输出的作用。GRU 的模型结构要较 LSTM 的简单, 而在同一组超参数下, GRU 的模型参数要明显少于 LSTM, 故理论上在处理较小规模的序列数据时更有优势。

从上面的数学模型中我们不难发现, 时刻 t 的输出只受时刻 t 及以前的输入影

响，不过在一些场景下我们会考虑时刻 t 以后的输入是会不会与时刻 t 的输出有关。譬如在文本意图识别时，英语有一种后置定语语法，在句子中一个单词 A 由出现在它后面的一组单词来修饰，那么为了正确理解在时刻 t 出现的单词 A ，我们就需要引入时刻 t 以后的信息。为此 Schuster 和 Paliwal^[67] 提出了双向递归神经网络（Bidirectional Recurrent Neural Network, BRNN），如图3.6所示，一个双向递归神经网络包含两个 RNN，对于一组序列数据，分别以其原本方向和反方向输入到两个 RNN 中，从而获得两个方向上与 t 时刻相关的信息。根据此思想，RNN 可以替换成任何具体的递归神经网络如 LSTM 和 GRU。

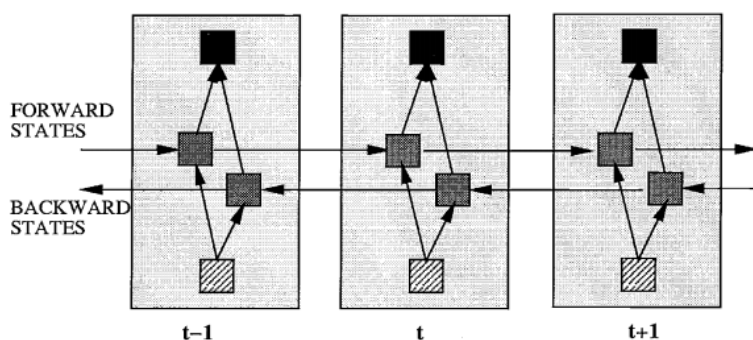


图 3.6 双向递归神经网络，引自^[67]

3.4.2.3 注意力机制

注意力机制（Attention Mechanism）受启发于人类的视觉处理机制，当人们尝试理解一个画面的内容时，他们不会马上对画面里的所有细节都进行处理，而是首先根据需要找出相关信息所在的部位，再对该部分的内容做进一步处理。当人们对同类型画面有一定观察后，会总结出关键信息出现的模式，并在处理新的画面时优先采用这些模式来理解其中的内容。Mnih 等人^[68] 就提出了结合递归神经网络和注意力机制来实现对手写数字的识别，不是透过预处理而是直接依赖注意力机制找出图片中数字所在的位置。其后 Bahdanau 等人^[69] 同样尝试了在递归神经网络中加入注意力机制，但这次的应用场景从计算机视觉转移到了自然语言处理的机器翻译，其注意力机制的模型如下：

$$e_{ij} = a(s_{i-1}, h_j) \quad (3-11)$$

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{T_x} \exp(e_{ik})} \quad (3-12)$$

$$c_i = \sum_{j=1}^{T_x} \alpha_{ij} h_j \quad (3-13)$$

其中 s_{i-1} 表示上一时刻 ($t-1$) 的隐藏状态, h_j 表示在时刻 j 的输入信息, a 是一个函数以计算 s_{i-1} 和 h_j 关联性, $\alpha_{ij} \in (0, 1]$ 表示对输出时刻 i , 给予输入时刻 j 的信息的注意力, 最后根据各个时刻应该给予的注意力结合所有输入信息, 得出用于时刻 i 输出的浓缩信息 c_i 。其后 Vaswani 等人^[70] 则仅使用注意力机制搭建的人工神经模型在机器翻译领域达到了引人注目的性能。虽然在近年不同研究^{[71][72][73]} 中同现了各种注意力机制的变型, 但这些模型可以统一描述如下:

$$\alpha_{ij} = a(q_i, k_j) \quad (3-14)$$

$$c_i = \frac{\sum_{j=1}^{T_x} \alpha_{ij} v_j}{\sum_{j=1}^{T_x} \alpha_{ij}} \quad (3-15)$$

其中 q_i 表示输出时刻 i 对应的一个请求, k_j 和 v_j 表示输入时刻 j 的一个键值对, a 是一个数计算请求和键之间关联性, 以计算对于输出时刻 i 应给予输入时刻 j 的信息 v_j 的注意力 α_{ij} , 最后以加权平均计算出输出时刻 i 相关的背景信息 c_i 。一些研究^{[20][21]} 则针对其中的隐藏状态 α_{ij} 分析注意力机制的具体运行, 并得出了和人工理解一致的结果。

3.4.3 集成学习

集成学习 (Ensemble Learning) 的目的在于结合多个模型以达到比其中任何一个模型都要好的识别性能。集成学习方法包含了多个类别, 其中常用于分类问题的是基于投票的方法, 譬如多数投票 (Majority Voting)、加权多数投票 (Weighted Majority Vote)、加权平均概率投票 (Soft Voting)。假设对于 N 分类问题, 有 T 个基础模型 $h_i, i \in [1, T]$, 对于输入数据 x , $h_i(x) \in \{0, 1\}^N$ 为各个模型的分类结果。对于多数投票其最终预测结果 $H_M(x)$ 如下:

$$H_M(x) = \frac{1}{T} \sum_{i=1}^T h_i(x) \quad (3-16)$$

对于加权多数投票, 其最终预测结果 $H_{WM}(x)$ 如下:

$$H_{WM}(x) = \sum_{i=1}^T w_i h_i(x) \quad (3-17)$$

其中 $\sum_{i=1}^T w_i = 1, w_i \geq 0$ 为每个模型的权重，相当于假设不同模型具有不同的可信度。而加权平均概率投票要求模型能给出样本在不同类别上的概率分布 $p_i(x) \in [0, 1]^N$ ，其最终预测概率分布 $P_S(x)$ 如下：

$$P_S(x) = \sum_{i=1}^T w_i p_i(x) \quad (3-18)$$

其中 $\sum_{i=1}^T w_i = 1, w_i \geq 0$ ，和加权多数投票中 w_i 的原理相同，最后取概率最高的类别作为最终分类结果。

3.5 本章小结

在本章中，我们对研究框架中关键模块相关的三个技术领域进行了介绍。对文本预处理，我们针对社交媒体上的英语文本给出了多个预处理的步骤，不过对于具体的数据集，建议由人工检查语料中出现频率较高的模式，并根据研究话题的需要进行特殊处理。对于特征提取，我们详细描述了词嵌入的算法和要注意的事项，另外简单介绍如何从词汇、句法、语义三个方面提取特征。最后我们介绍了意图识别中常用到的三个类别的模型，首先是传统机器学习和人工神经网络，分别对他们的核心思想和数学模型给出了简要的说明，最后介绍了集成学习中基于投票方法，它将用于我们的集成识别系统的设计当中。

第 4 章 面向微博的反讽识别

4.1 本章引论

pass

4.2 本章小结

pass

第 5 章 面向三轮对话的情感识别

5.1 本章引论

pass

5.2 本章小结

pass

插图索引

图 1.1	Plutchik ^[4] 提出的情感轮模型	2
图 1.2	Hugo ^[11] 的情感立方体模型	3
图 1.3	引入多个领域信息的反讽识别神经网络模型, 引自 ^[28]	7
图 1.4	基于卷积神经网络的微博情感分类模型, 引自 ^[35]	8
图 2.1	意图识别的研究框架	12
图 3.1	Word2vec 提供的两个算法的模型, 引自 ^[48]	18
图 3.2	支持向量机在处理二类问题的示意图。其中黑点和白点分别对应两个类别的样本, L_1 不能对两类样本作区分, L_2 和 L_3 均成功区分两个类别的样本, 但 L_3 和两个类别样本的距离更大, 为更优解	20
图 3.3	生物的神经元与感知器的模型示意图, 引自 ^[56]	21
图 3.4	卷积神经网络模型	22
图 3.5	长短时记忆模型, 引自 ^[65]	24
图 3.6	双向递归神经网络, 引自 ^[67]	25

表格索引

公式索引

公式 3-1	23
公式 3-2	23
公式 3-3	23
公式 3-4	23
公式 3-5	23
公式 3-6	23
公式 3-7	23
公式 3-8	24
公式 3-9	24
公式 3-10	24
公式 3-11	25
公式 3-12	25
公式 3-13	26
公式 3-14	26
公式 3-15	26
公式 3-16	26
公式 3-17	27
公式 3-18	27

参考文献

- [1] Banerjee S, Dutta U. Detection of emotions in text: A survey: volume 03[M]. [S.l.: s.n.], 2015
- [2] Chatterjee A, Narahari K N, Joshi M, et al. Semeval-2019 task 3: Emocontext: Contextual emotion detection in text[C]//Proceedings of The 13th International Workshop on Semantic Evaluation (SemEval-2019). Minneapolis, Minnesota: [s.n.], 2019.
- [3] Alm C O, Dan R, Sproat R. Emotions from text: Machine learning for text-based emotion prediction[C]//Conference on Hlt/emnlp. [S.l.: s.n.], 2005.
- [4] Plutchik R, Kellerman H. Emotion: Theory, research and experience[J]. In Theories of emotion, 1980, 11: 399.
- [5] Ekman P. An argument for basic emotions[J]. Cognition & Emotion, 1992, 6(3-4): 169-200.
- [6] Russell J. A circumplex model of affect[J]. Journal of Personality and Social Psychology, 1980, 39: 1161-1178.
- [7] M. Bradley M, Greenwald M, C. Petry M, et al. Remembering pictures: Pleasure and arousal in memory[J]. Journal of experimental psychology. Learning, memory, and cognition, 1992, 18: 379-90.
- [8] Watson D, Tellegen A. Toward a consensual structure of mood[J]. Psychological bulletin, 1985, 98: 219-35.
- [9] Rubin D C, Talarico J M. A comparison of dimensional models of emotion: evidence from emotions, prototypical events, autobiographical memories, and words[J]. Memory, 2009, 17(8): 802-808.
- [10] Mehrabian A. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament[J]. Current Psychology, 1996, 14(4): 261-292.
- [11] Hugo L. A new three-dimensional model for emotions and monoamine neurotransmitters[J]. Medical Hypotheses, 2012, 78(2): 341-348.
- [12] Mohammad S, Bravo-Marquez F, Salameh M, et al. Semeval-2018 task 1: Affect in tweets[C]//Proceedings of The 12th International Workshop on Semantic Evaluation. [S.l.: s.n.], 2018: 1-17.
- [13] Turney P D. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews[C]//Proceedings of the 40th annual meeting on association for computational linguistics. [S.l.]: Association for Computational Linguistics, 2002: 417-424.
- [14] Pang B, Lee L. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts[C]//Proceedings of the 42nd annual meeting on Association for Computational Linguistics. [S.l.]: Association for Computational Linguistics, 2004: 271.
- [15] 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): volume 1. [S.l.: s.n.], 2015: 1014-1023.

- [16] Khan A, Baharudin B, Khan K. Sentiment classification using sentence-level lexical based[J]. Trends in Applied Sciences Research, 2011, 6(10): 1141-1157.
- [17] Li Y, Lin Y, Zhang J, et al. Constructing domain-dependent sentiment lexicons automatically for sentiment analysis[J]. Information Technology Journal, 2013, 12(5): 990-996.
- [18] Dos Santos C, Gatti M. Deep convolutional neural networks for sentiment analysis of short texts[C]//Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers. [S.l.: s.n.], 2014: 69-78.
- [19] Che W, Zhao Y, Guo H, et al. Sentence compression for aspect-based sentiment analysis[J]. IEEE/ACM transactions on audio, speech, and language processing, 2015, 23(12): 2111-2124.
- [20] Wang Y, Huang M, Zhao L, et al. Attention-based lstm for aspect-level sentiment classification [C]//Proceedings of the 2016 conference on empirical methods in natural language processing. [S.l.: s.n.], 2016: 606-615.
- [21] Pontiki M, Galanis D, Pavlopoulos J, et al. Semeval-2014 task 4: Aspect based sentiment analysis [C]//Proceedings of The 8th International Workshop on Semantic Evaluation (SemEval-2014). [S.l.: s.n.], 2014: 27-35.
- [22] Tang D, Qin B, Liu T. Aspect level sentiment classification with deep memory network[J]. arXiv preprint arXiv:1605.08900, 2016.
- [23] Tsur O, Davidov D, Rappoport A. Icwsm—a great catchy name: Semi-supervised recognition of sarcastic sentences in online product reviews[C]//Fourth International AAAI Conference on Weblogs and Social Media. [S.l.: s.n.], 2010.
- [24] Davidov D, Tsur O, Rappoport A. Semi-supervised recognition of sarcastic sentences in twitter and amazon[C]//Proceedings of the fourteenth conference on computational natural language learning. [S.l.: Association for Computational Linguistics, 2010: 107-116.
- [25] Reyes A, Rosso P, Veale T. A multidimensional approach for detecting irony in twitter[J]. Language resources and evaluation, 2013, 47(1): 239-268.
- [26] Kunneman F, Liebrecht C, Van Mulken M, et al. Signaling sarcasm: From hyperbole to hashtag [J]. Information Processing & Management, 2015, 51(4): 500-509.
- [27] Littlestone N. Learning quickly when irrelevant attributes abound: A new linear-threshold algorithm[J]. Machine learning, 1988, 2(4): 285-318.
- [28] Poria S, Cambria E, Hazarika D, et al. A deeper look into sarcastic tweets using deep convolutional neural networks[J]. arXiv preprint arXiv:1610.08815, 2016.
- [29] Madhusudhanan S, Moorthi D M. A survey on sentiment analysis[J]. Indian Journal of Computer Science and Engineering, 2018, 9(2).
- [30] 谢丽星, 周明, 孙茂松. 基于层次结构的多策略中文微博情感分析和特征抽取[J]. 中文信息学报, 2012, 26(1): 73-84.
- [31] Go A, Huang L, Bhayani R. Twitter sentiment analysis[J]. Entropy, 2009, 17: 252.
- [32] Paltoglou G, Thelwall M. Twitter, myspace, digg: Unsupervised sentiment analysis in social media[J]. ACM Transactions on Intelligent Systems and Technology (TIST), 2012, 3(4): 66.
- [33] Khan FH, Bashir S, Qamar U. Tom: Twitter opinion mining framework using hybrid classification scheme[J]. Decision Support Systems, 2014, 57: 245-257.

-
- [34] Angiani G, Ferrari L, Fontanini T, et al. A comparison between preprocessing techniques for sentiment analysis in twitter.[C]//KDWeb. [S.l.: s.n.], 2016.
 - [35] 张海涛, 王丹, 徐海玲, 等. 基于卷积神经网络的微博舆情情感分类研究[J]. 情报学报, 2018, 37: 695-702.
 - [36] Van Hee C, Lefever E, Hoste V. Semeval-2018 task 3: Irony detection in english tweets[C]// Proceedings of The 12th International Workshop on Semantic Evaluation (SemEval-2018). [S.l.: s.n.], 2018: 39-50.
 - [37] 刘丹丹, 邱恒清, 赵应丁. 基于 svm 的中文微博情感识别与分类研究[J]. 中国新通信, 2015, 17(21): 48-51.
 - [38] 邓钊, 贾修一, 陈家骏. 面向微博的中文反语识别研究[J]. 计算机工程与科学, 2015, 37(12): 2312-2317.
 - [39] Zahiri S M, Choi J D. Emotion detection on tv show transcripts with sequence-based convolutional neural networks[Z]. [S.l.: s.n.], 2017.
 - [40] Jackson P, Moulinier I. Natural language processing for online applications[M]. [S.l.: John Benjamins, 2007
 - [41] Mitkov R. The oxford handbook of computational linguistics[M]. [S.l.: Oxford University Press, 2004
 - [42] Ahmed F, Luca E W D, Nürnberger A. Revised n-gram based automatic spelling correction tool to improve retrieval effectiveness[J]. Polibits, 2009(40): 39-48.
 - [43] Nejja M, Yousfi A. The context in automatic spell correction[J]. Procedia Computer Science, 2015, 73: 109-114.
 - [44] Baziotis C, Pelekis N, Doukeridis C. Datastories at semeval-2017 task 4: Deep lstm with attention for message-level and topic-based sentiment analysis[C]//Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017). Vancouver, Canada: Association for Computational Linguistics, 2017: 747-754.
 - [45] Kouloumpis E, Wilson T, Moore J. Twitter sentiment analysis: The good the bad and the omg! [C]//Fifth International AAAI conference on weblogs and social media. [S.l.: s.n.], 2011.
 - [46] Joshi A, Sharma V, Bhattacharyya P. Harnessing context incongruity for sarcasm detection[C]// Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers): volume 2. [S.l.: s.n.], 2015: 757-762.
 - [47] Baziotis C, Athanasiou N, Papalampidi P, et al. Ntua-slp at semeval-2018 task 3: Tracking ironic tweets using ensembles of word and character level attentive rnns[J]. arXiv preprint arXiv:1804.06659, 2018.
 - [48] Mikolov T, Chen K, Corrado G, et al. Efficient estimation of word representations in vector space [J]. arXiv preprint arXiv:1301.3781, 2013.
 - [49] Deerwester S, Dumais S T, Furnas G W, et al. Indexing by latent semantic analysis[J]. Journal of the American society for information science, 1990, 41(6): 391-407.
 - [50] Pennington J, Socher R, Manning C D. Glove: Global vectors for word representation[C/OL]// Empirical Methods in Natural Language Processing (EMNLP). 2014: 1532-1543. <http://www.aclweb.org/anthology/D14-1162>.

- [51] Cortes C, Vapnik V. Support-vector networks[J]. Machine learning, 1995, 20(3): 273-297.
- [52] Drucker H, Burges C J, Kaufman L, et al. Support vector regression machines[C]//Advances in neural information processing systems. [S.l.: s.n.], 1997: 155-161.
- [53] Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain.[J]. Psychological review, 1958, 65(6): 386.
- [54] Rumelhart D E, Hinton G E, Williams R J. Learning internal representations by error propagation [R]. [S.l.]: California Univ San Diego La Jolla Inst for Cognitive Science, 1985.
- [55] Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets[J]. Neural computation, 2006, 18(7): 1527-1554.
- [56] Basheer I A, Hajmeer M. Artificial neural networks: fundamentals, computing, design, and application[J]. Journal of microbiological methods, 2000, 43(1): 3-31.
- [57] LeCun Y, Boser B, Denker J S, et al. Backpropagation applied to handwritten zip code recognition [J]. Neural computation, 1989, 1(4): 541-551.
- [58] Waibel A, Hanazawa T, Hinton G, et al. Phoneme recognition using time-delay neural networks [J]. Backpropagation: Theory, Architectures and Applications, 1995: 35-61.
- [59] Hubel D H, Wiesel T N. Receptive fields of single neurones in the cat's striate cortex[J]. The Journal of physiology, 1959, 148(3): 574-591.
- [60] Hubel D H, Wiesel T N. Receptive fields and functional architecture of monkey striate cortex[J]. The Journal of physiology, 1968, 195(1): 215-243.
- [61] Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position[J]. Biological cybernetics, 1980, 36(4): 193-202.
- [62] Zhang X, Zhao J, LeCun Y. Character-level convolutional networks for text classification[C]//Advances in neural information processing systems. [S.l.: s.n.], 2015: 649-657.
- [63] Elman J L. Finding structure in time[J]. Cognitive science, 1990, 14(2): 179-211.
- [64] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural computation, 1997, 9(8): 1735-1780.
- [65] Graves A, Jaitly N, Mohamed A r. Hybrid speech recognition with deep bidirectional lstm[C]//2013 IEEE workshop on automatic speech recognition and understanding. [S.l.]: IEEE, 2013: 273-278.
- [66] Cho K, Van Merriënboer B, Gulcehre C, et al. Learning phrase representations using rnn encoder-decoder for statistical machine translation[J]. arXiv preprint arXiv:1406.1078, 2014.
- [67] Schuster M, Paliwal K K. Bidirectional recurrent neural networks[J]. IEEE Transactions on Signal Processing, 1997, 45(11): 2673-2681.
- [68] Mnih V, Heess N, Graves A, et al. Recurrent models of visual attention[C]//Advances in neural information processing systems. [S.l.: s.n.], 2014: 2204-2212.
- [69] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate[J]. arXiv preprint arXiv:1409.0473, 2014.
- [70] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//Advances in Neural Information Processing Systems. [S.l.: s.n.], 2017: 5998-6008.
- [71] Luong M T, Pham H, Manning C D. Effective approaches to attention-based neural machine translation[J]. arXiv preprint arXiv:1508.04025, 2015.

- [72] Graves A, Wayne G, Danihelka I. Neural turing machines[J]. arXiv preprint arXiv:1410.5401, 2014.
- [73] Xu K, Ba J, Kiros R, et al. Show, attend and tell: Neural image caption generation with visual attention[C]//International conference on machine learning. [S.l.: s.n.], 2015: 2048-2057.
- [74] 薛瑞尼. THUTHESIS: 清华大学学位论文模板[EB/OL]. 2017. <https://github.com/xueruini/thuthesis>.

致 谢

衷心感谢导师 xxx 教授和物理系 xxx 副教授对本人的精心指导。他们的言传身教将使我终生受益。

在美国麻省理工学院化学系进行九个月的合作研究期间，承蒙 xxx 教授热心指导与帮助，不胜感激。感谢 xx 实验室主任 xx 教授，以及实验室全体老师和同学们的热情帮助和支持！本课题承蒙国家自然科学基金资助，特此致谢。

感谢 L^AT_EX 和 THU^THESIS^[74]，帮我节省了不少时间。

声 明

本人郑重声明：所呈交的学位论文，是本人在导师指导下，独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本学位论文的研究成果不包含任何他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。

签 名：_____ 日 期：_____

个人简历、在学期间发表的学术论文与研究成果

个人简历

xxxx 年 xx 月 xx 日出生于 xx 省 xx 县。

xxxx 年 9 月考入 xx 大学 xx 系 xx 专业, xxxx 年 7 月本科毕业并获得 xx 学士学位。

xxxx 年 9 月免试进入 xx 大学 xx 系攻读 xx 学位至今。

发表的学术论文

- [1] Yang Y, Ren T L, Zhang L T, et al. Miniature microphone with silicon- based ferroelectric thin films. Integrated Ferroelectrics, 2003, 52:229-235. (SCI 收录, 检索号:758FZ.)
- [2] 杨轶, 张宁欣, 任天令, 等. 硅基铁电微声学器件中薄膜残余应力的研究. 中国机械工程, 2005, 16(14):1289-1291. (EI 收录, 检索号:0534931 2907.)
- [3] 杨轶, 张宁欣, 任天令, 等. 集成铁电器件中的关键工艺研究. 仪器仪表学报, 2003, 24(S4):192-193. (EI 源刊.)
- [4] Yang Y, Ren T L, Zhu Y P, et al. PMUTs for handwriting recognition. In press. (已被 Integrated Ferroelectrics 录用. SCI 源刊.)
- [5] Wu X M, Yang Y, Cai J, et al. Measurements of ferroelectric MEMS microphones. Integrated Ferroelectrics, 2005, 69:417-429. (SCI 收录, 检索号:896KM)
- [6] 贾泽, 杨轶, 陈兢, 等. 用于压电和电容麦克风的体硅腐蚀相关研究. 压电与声光, 2006, 28(1):117-119. (EI 收录, 检索号:06129773469)
- [7] 伍晓明, 杨轶, 张宁欣, 等. 基于 MEMS 技术的集成铁电硅微麦克风. 中国集成电路, 2003, 53:59-61.

研究成果

- [1] 任天令, 杨轶, 朱一平, 等. 硅基铁电微声学传感器畴极化区域控制和电极连接的方法: 中国, CN1602118A. (中国专利公开号)

- [2] Ren T L, Yang Y, Zhu Y P, et al. Piezoelectric micro acoustic sensor based on ferroelectric materials: USA, No.11/215, 102. (美国发明专利申请号)

综合论文训练记录表

学生姓名		学号		班级	
论文题目					
主要内容以及进度安排	<div>指导教师签字：_____</div> <div>考核组组长签字：_____</div> <div>年 月 日</div>				
中期考核意见	<div>考核组组长签字：_____</div> <div>年 月 日</div>				

指导教师评语	<div>指导教师签字：_____</div> <div>年 月 日</div>
评阅教师评语	<div>评阅教师签字：_____</div> <div>年 月 日</div>
答辩小组评语	<div>答辩小组组长签字：_____</div> <div>年 月 日</div>

总成绩：_____

教学负责人签字：_____

年 月 日