文章编号：2096-1472(2022)-09-01-03
DOI:10.19644/j.cnki.issn2096-1472.2022.009.001

# 结合显式句法依赖与分层注意力进行方面级情感分析

范明炜，张云华

(浙江理工大学信息学院，浙江 杭州 310018)
✉719068852@qq.com; 605498519@qq.com

**摘　要**：针对方面级情感分析中未能充分利用显式句法依赖的问题，提出基于语法依赖的分层注意力网络。对每个单词与方面词之间的语法路径进行建模，表征每个词对方面词的句法表示，将生成的句法表示反馈到关注层来推断权重。通过分层注意力对单词和句子赋予不同的注意力权重，多方面帮助模型增加对重要部分的注意力。实验结果表明，该方法在SemEval−2014中优于现有的算法。

**关键词**：情感分析；句法依赖；注意力机制
**中图分类号**：TP391.1　**文献标识码**：A



# Aspect-Level Sentiment Analysis Combining Explicit Syntactic Dependencies and Hierarchical Attention

FAN Mingwei, ZHANG Yunhua

( *School of Information Science and Technology*, *Zhejiang Sci-Tech University*, *Hangzhou* 310018, *China*)
✉719068852@qq.com; 605498519@qq.com

**Abstract:** Aiming at the problem of underutilization of explicit syntactic dependencies in aspect-level sentiment analysis, this paper proposes a hierarchical attention network based on grammar dependencies. The syntax path between each word and aspect word is modeled to represent the syntactic representation of each word to the aspect word, and the generated syntactic representation is fed back to the attention layer to infer the weight. Through hierarchical attention, different attention weights are assigned to words and sentences, which helps the model to pay more attention to important parts in many ways. Experimental results show that this method is superior to existing algorithms in SemEval-2014.

**Keywords:** sentiment analysis; syntactic dependency; attention mechanism

## 1 引言(Introduction)

方面级这种细粒度的情感分析[1−2]解决了针对一段评论的不同方面，情感的判断可能出现两种相反结果的问题，因此目前方面级情感分析逐渐成为研究的热点，对商品评论、推荐系统等领域具有重要意义。

现有的深度学习模型在情感分析方面取得了较好的效果。基于语义的方法将输入的句子看作单词序列，通过注意力建模，例如RNN、Transformer等[3]；基于语法的方法通过引入句法依赖关系树构造输入句子的语法结构，采用GNN通过依赖关系树上下文词的表示来丰富方面表示[4]；但它们均未充分利用上下文词与方面词之间的语法依赖[5]。XIONG等[6]采用字符级别词嵌入实现文本分类。YANG等做出改进，提出了基于TD−LSTM[7]的方法。但邵兴林通过实体信息与属性信息的比较，发现实体信息更加重要，同时评论可能会有多个句子，选出情感更加强烈的句子也具有重要意义[8]。

受此启发，本文提出结合分层注意力机制[9]与显式句法依赖的多层网络，结合依赖路径编码和实体信息的嵌入表示一并发送到注意力层。构建单词−句子、句子−文档的层次结构，通过加入多级注意力机制，使模型对不同单词、句子赋予不同的注意力权重，对最后的结果做出更准确的判断。

## 2 相关工作(Related work)

### 2.1 Stanford Parser

Stanford Parser[10]是基于概率统计的开源句法分析器。基于Penn Treebank作为分析器的训练数据，面向中文、英文等

语种提供句法分析功能，可以输出句法分析树，如图1所示。

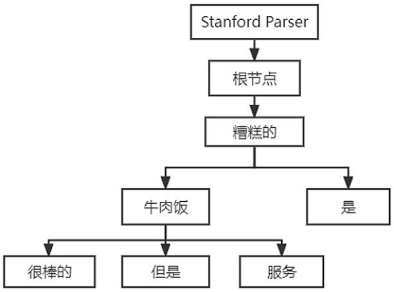


图1 句法分析树

Fig.1 Syntactic parse tree

### 2.2 层次注意力

注意力机制就是关注输入权重分配，可以理解成一个由查询矩阵$Q$和对应的键$K$，以及需加权平均的值$V$构成的一层感知机。

层次注意力用于解决多层次问题[11]。如在分析评论时，把词作为一层，把句子作为一层，这样就有了多层，下一层对上一层产生影响，因此建立了一种堆叠的层次注意力模型，如图2所示。

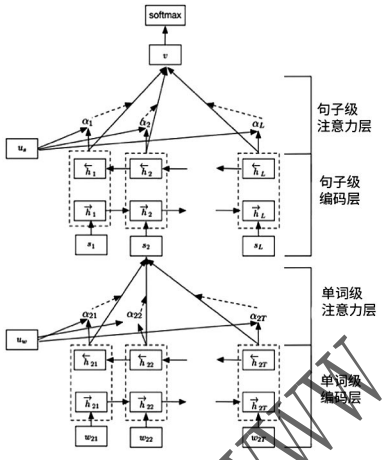


图2 层次注意力模型

Fig.2 Hierarchical attention model

### 2.3 LSTM

LSTM是一种特殊的RNN，主要是为了解决长序列训练过程的梯度消失和爆炸问题，长序列使用LSTM有更好的表现[12]。传输过程中，通过门控状态来控制需要长时间记忆的和忘记不重要的信息，其内部可划分为忘记阶段、选择记忆阶段以及输出阶段，如图3所示。

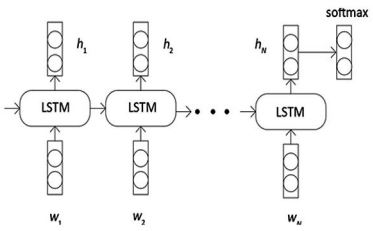


图3 LSTM

Fig.3 LSTM

## 3 模型构建(Model building)

给定长度为$n$的句子$s=\{w1, w2, w3, \cdots, wn\}$，以及长度为$m(0<m<n)$的方面项$a=\{w[i+1], \cdots, w[i+m]\}$，其中$a \in s$，$a$可以是词或者短语。目标是对方面项$a$进行情感分析，最终分为积极、消极、中性。

图4展示了模型的整体设计。在词级别，通过句法分析获取句法依赖树，从而得到每个单词到方面项的路径编码和距离。同时获取句子中单词的词性及方面项的实体信息，结合上述路径编码和距离一起馈送到各自的编码层，通过注意力网络得到句子表示。在句子级别，评论中的其他句子表示获取同上，通过LSTM及注意力给予重要的句子更大的权重。这样，单词和句子都有了自己的注意力权重。

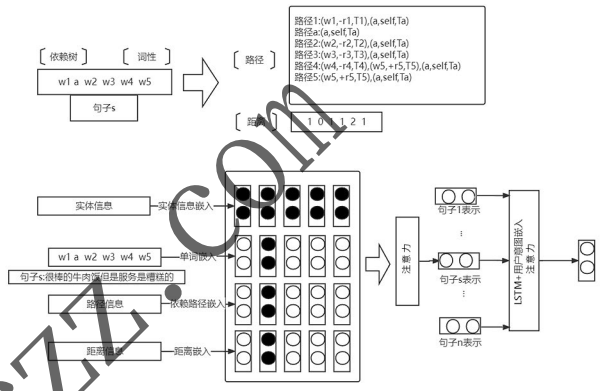


图4 模型整体设计

Fig.4 Overall design of the model

### 3.1 输入层

#### 3.1.1 句法依赖分析和词性

通过Stanford Parser获取句子结构树与每个单词的词性。词性对于情感分析任务是非常重要的，通常形容词和动词比名词表达的程度更深。句法结构树包含词对的关系，在句法依赖路径编码中需要用到。

#### 3.1.2 句法依赖路径与距离

句法依赖树中每个词到方面词的有向关系路径称为句法依赖路径，路径上边的个数称为距离。例如图1中的依赖树，针对“牛肉饭”这一方面词，句中单词到它的距离为{1，0，1，1，2，1}。针对“很棒的”到“服务”这一方面词的路径，“很棒的—(−amod)→牛肉饭—(+conj)→服务”可以被转换为“[(很棒的，−amod，JJ)，(牛肉饭，+conj，NN)，(服务，self，NN)]”。“+”和“−”表示有向路径的正反，即关系的方向，并且在最后为方面词附加一个预定义的关系“self”。简单地将依赖路径进行分解容易丢失路径的全局特征，因此使用LSTM对依赖路径序列进行编码，如图5所示。

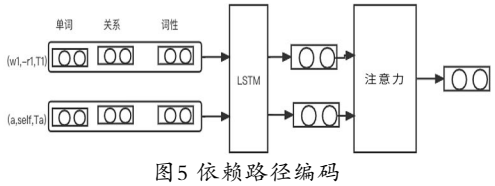


图5 依赖路径编码

Fig.5 Dependent path encoding

#### 3.1.3 句子向量表示

实体信息对方面级情感分析具有重要作用，如“牛肉饭”这一方面词的实体信息为“食物”，将“食物”通过词向量矩阵获取词向量，同时结合依赖路径编码与距离编码，将上述组成的向量送入注意力层，获得句子向量表示。

### 3.2 分层注意力

一段评论可能包含多个句子，给句子加入注意力可以使模型提取更适合的句子进行情感极性分析。本文加入新一层句子级别的注意力机制，依照句子向量表示的方法，对每一句话都获取向量表示，将这些句子向量表示作为LSTM新的输入，再将LSTM的输出与实体信息结合，通过注意力机制提取对评论情感极性更加重要的句子，以获取最终评论的向量表示。

### 3.3 模型训练

利用交叉熵损失函数和L2正则化对模型进行训练，公式如下：

$$Loss = -\sum_{i}\sum_{j} y_i^j \log \hat{y}_i^j + \lambda \|\theta\|$$

其中，$i$和$j$分别代表评论和类别的下标，$\lambda$是L2的正则系数，$\theta$代表参数集合。

## 4 实验结果分析(Experimental results and analysis)

### 4.1 数据集

为了验证模型的有效性，本文使用了SemEval 2014 Task 4中的restaurants数据集，包含三种情感和五个方面，如表1所示。

表1 数据集分类统计表

Tab.1 Statistical table of the dataset classification

| 情感分类 | 积极 | | 消极 | | 中性 | |
|---|---|---|---|---|---|---|
| | 训练集 | 测试集 | 训练集 | 测试集 | 训练集 | 测试集 |
| Food | 867 | 302 | 209 | 69 | 90 | 31 |
| Service | 324 | 101 | 218 | 63 | 20 | 3 |
| Price | 179 | 51 | 115 | 28 | 10 | 1 |
| Ambience | 263 | 76 | 98 | 21 | 23 | 8 |
| Anecdotes | 546 | 127 | 199 | 41 | 357 | 51 |
| 总计 | 2,179 | 657 | 839 | 222 | 500 | 94 |

### 4.2 注意力机制有效性

对于例句“很棒的牛肉饭，但是服务是糟糕的”，当给定实体为“食物”时，“牛肉饭”和“很棒的”被赋予了更高的权重，在句子情感分析中二者起到重要作用。对于该评论“很棒的牛肉饭”和“但是服务是糟糕的”这两个句子，第一句在情感分析中起到重要的作用，而这句话是实体“食物”对应的句子，结果符合预期，如图6所示。

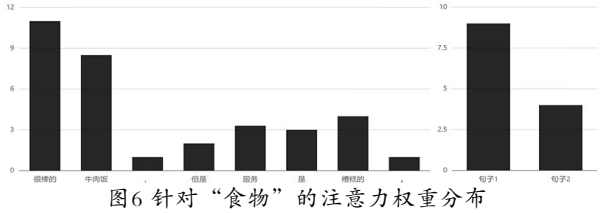


图6 针对“食物”的注意力权重分布

Fig.6 Attention weight distribution for "food"

同理，将实体换为“服务”时，单词和句子的注意力权重数据也同样符合预期，如图7所示。

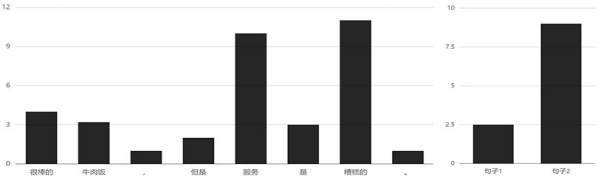


图7 针对“服务”的注意力权重分布

Fig.7 Attention weight distribution for "service"

### 4.3 模型对比

通过在相同数据集中与其他模型进行对比，验证本文模型的有效性，主要采用F1值和Acc值(准确率)进行评估。对比的模型主要有BiLSTM-AMM、BiGRU-AAM、Bi-LSTM、ATAE-LSTM等。

如表2所示，从实验结果来看，本文的方法在数据集上相比于基本的深度学习模型，Acc值和F1值都有所提高，主要是因为模型通过引入实体信息，充分利用显式语法结构获取到更加有用的信息；其次，分层注意力机制的引入也使得结果变得更加精确。

表2 不同模型在数据集中的结果

Tab.2 Results of different models on the dataset

| 模型 | 指标 | 情感极性 | | |
|---|---|---|---|---|
| | | 积极/% | 消极/% | 中性/% |
| BiLSTM-AMM | Acc | 89.81 | 81.52 | 56.32 |
| | F1 | 89.95 | 81.62 | 54.05 |
| BiGRU-AAM | Acc | 90.15 | 83.72 | 58.21 |
| | F1 | 90.05 | 82.39 | 55.29 |
| Bi-LSTM | Acc | 90.02 | 80.24 | 36.52 |
| | F1 | 86.99 | 78.90 | 43.16 |
| ATAE-LSTM | Acc | 91.21 | 81.58 | 42.13 |
| | F1 | 88.85 | 79.30 | 49.37 |
| 本文方法 | Acc | 91.30 | 84.15 | 59.37 |
| | F1 | 91.16 | 83.44 | 56.38 |

## 5 结论(Conclusion)

针对方面级情感分析中未对句法结构信息与属性信息进行深度挖掘的问题，本文提出的模型一方面利用句法结构、实体信息加强特征获取的能力，另一方面利用分层注意力机制使模型能够赋予重要单词和句子更大的权重。从实验结果来看，该模型能有效提高情感分类的效果。

## 参考文献(References)

[1] 李林川.方面级文本情感分析的研究与应用[D].兰州:兰州交通大学,2021.

[2] 王海燕,陶皖,余玲艳,等.文本细粒度情感分析综述[J].河南科技学院学报(自然科学版),2021,49(04):67-76.
[3] 田元,周晓蕾,周幂,等.学习情感分析方法研究综述[J].中国教育信息化,2021(22):1-6.
[4] 张合桥,苟刚,陈青梅.基于图神经网络的方面级情感分析[J].计算机应用研究,2021,38(12):3574-3580.
[5] KE W J, GAO J H, SHEN H W, et al. Incorporating explicit syntactic dependency for aspect level sentiment classification[J]. Neurocomputing, 2021, 456:394-406.
[6] XIONG S F, LV H L, ZHAO W T, et al. Towards twitter sentiment classification by multi-level sentiment-enriched word embeddings[J]. Neurocomputing, 2018, 275:2459-2466.
[7] YANG M, QU Q, CHEN X J, et al. Feature-enhanced attention network for target-dependent sentiment classification[J]. Neurocomputing, 2018, 307:91-97.
[8] 邵兴林.基于深度学习的细粒度情感分析[D].北京:北京邮电大学,2020.
[9] 王彦卿.基于BERT和分层注意力网络的方面级情感分析[D].哈尔滨:哈尔滨理工大学,2021.
[10] 项炜,金澎.大规模语料库上的Stanford和Berkeley句法分析器性能对比分析[J].电脑知识与技术,2013,9(08):1984-1986.
[11] 余本功,朱晓洁,张子薇.基于多层次特征提取的胶囊网络文本分类研究[J].数据分析与知识发现,2021,5(06):93-102.
[12] 文万志,姜文轩,葛威,等.一种基于深度学习的实体消歧技术[J].南通大学学报(自然科学版),2021,20(04):23-30.

**作者简介：**
范明炜(1996-)，男，硕士生.研究领域：软件工程技术.
张云华(1965-)，男，博士，研究员.研究领域：软件工程，系统仿真，智能信息处理.

(上接第21页)

漏。系统分配给应用程序的内存资源是有限的，当应用程序向系统申请内存时，系统会在堆内存中开辟一块内存空间，如果应用程序中的全局变量、自定义事件没有被回收和销毁，久而久之，内存就会被占满，发生内存泄漏的现象，导致软件卡顿直至崩溃。

(5)对于首屏加载过慢的问题或页面组件过多难以一次性加载的情况，采用异步组件的方式，只有用户触发了该组件，才会对该组件进行渲染。对加载过的组件同样进行缓存处理，避免二次加载造成资源浪费。

## 4 结论(Conclusion)

本文以药房药库管理系统为例，运用前端工程化思想，采用前后端分离架构，前端采用了Vue.js，以组件和模块的方式对视图部分进行开发；后端采用了Node.js，负责处理业务逻辑，提供数据接口，并结合MVVM模式、组件化、模块化等解决方案，达到组件间高效、协作、复用的效果，模块间互不影响，视图与数据分离，细化了开发者的分工协作，并从用户隐私安全、接口安全、系统安全三大角度进行了分析。提出组件状态的管理机制以及错误信息的控制措施，从整体上提升了项目可维护性和拓展性，同时提高了开发效率，降低了开发成本。解决了传统医院信息系统分层开发方式的系统性能差、维护难、开发效率低下、层级间协作不一致的问题，提高了医院的药品管理水平，促进了医院的信息化发展。

## 参考文献(References)

[1] 胡芳华.某三甲医院信息集成平台设计与应用[D].长沙:湖南大学,2017.
[2] 曾广海.基于Web前端组件化的个人博客系统的设计与实现[D].武汉:华中科技大学,2016.
[3] 杨江兵.基于物联网技术的现代化药房快速发药系统的研究[D].北京:北京邮电大学,2015.
[4] 汪彤.基于Node.js的图书共享平台的设计与实现[D].北京:北京邮电大学,2018.
[5] 胡佳静.基于electron的待办事项管理app开发[D].武汉:华中科技大学,2018.
[6] 李添译.基于MVVM中间件的航天企业综合业务管理系统的设计与实现[D].北京:北京邮电大学,2018.
[7] 杨磊.基于虚拟DOM的前端可视化编辑系统的设计与实现[D].武汉:华中科技大学,2017.
[8] 曹海歌.基于改进的Diff算法的Web前端性能优化及应用[D].武汉:华中师范大学,2016.
[9] 沈姝.NoSQL数据库技术及其应用研究[D].南京:南京信息工程大学,2012.
[10] WAKHARE S, PISE P, KHALATE R, et al. Secure login system using MD5 and AES attribute based encryption algorithm[J]. International Journal of Innovative Technology and Exploring Engineering, 2020, 9(8):810-814.
[11] 王晓强.基于HTML5的CSRF攻击与防御技术研究[D].秦皇岛:燕山大学,2013.

**作者简介：**
李思睿(2001-)，女，本科生.研究领域：人工智能.
郑大翔(1998-)，男，本科生.研究领域：计算机应用.
李志芳(1980-)，女，硕士，副教授.研究领域：人工智能，智慧医疗.本文通讯作者.