

# 基于条件随机场模型的汉语主谓短语自动识别研究

何一凡

西南交通大学人文学院中文系, 四川 成都

收稿日期: 2025年11月6日; 录用日期: 2025年12月2日; 发布日期: 2025年12月18日

## 摘要

主谓结构是汉语句子的核心语法骨架, 其精准识别是自然语言处理(NLP)领域语义解析、信息抽取等下游任务的关键前提。针对中文语言结构复杂性给主谓短语识别带来的挑战, 本文提出基于条件随机场(CRF)的汉语主谓短语自动识别方法, 以提升识别准确性与可靠性。研究以清华-汉语句法树库(TCT)为语料来源, 构建含39595个标注样本的数据集。预处理阶段, 通过自定义转换函数解决原始语料编码问题, 用正则表达式匹配主谓结构边界, 完成“起始(B) - 内部(I) - 其他(O)”标签标注, 并将数据格式化为“词语 - 词性 - 标签”三元组, 满足CRF模型训练需求。特征工程中, 设计word2features函数, 抓取词形及变形、词性、上下文前后1~2词及组合特征、句子边界(BOS/EOS)、词缀等多维度特征, 为模型提供支撑。采用sklearn-crfsuite库实现CRF, 以L-BFGS为优化算法, 设正则化系数( $C_1 = 0.2$ 、 $C_2 = 0.1$ )与最大迭代次数200次, 将数据集按9:1划分为训练集35635个样本与测试集3960个样本。实验显示, 模型加权F1值0.7459, I标签F1值0.7541, B标签F1值0.6739; 加权精确率0.7675、召回率0.7257, 模型对主谓结构内部成分识别较好, 但起始边界及长距离依赖识别需优化。研究证实, 词性与上下文组合特征可提升模型性能, 为汉语句法结构自动识别提供参考。

## 关键词

汉语主谓短语, 条件随机场(CRF), 自动识别, 特征工程

## A CRF-Based Approach for Automatic Identification of Chinese Subject-Predicate Phrases

Yifan He

Department of Chinese Language and Literature, School of Humanities, Southwest Jiaotong University, Chengdu Sichuan

Received: November 6, 2025; accepted: December 2, 2025; published: December 18, 2025

文章引用: 何一凡. 基于条件随机场模型的汉语主谓短语自动识别研究[J]. 现代语言学, 2025, 13(12): 349-356.

DOI: 10.12677/ml.2025.13121271

## Abstract

The subject-predicate structure constitutes the fundamental syntactic framework of Chinese sentences, and its precise identification is essential for various downstream natural language processing (NLP) tasks, including semantic parsing and information extraction. To address the challenges posed by the structural complexity of Chinese in recognizing subject-predicate phrases, this paper introduces an automated identification method based on Conditional Random Fields (CRF), aiming to improve both the accuracy and robustness of recognition. Using the Tsinghua Chinese Treebank (TCT) as the corpus, we constructed a dataset of 39,595 annotated samples. During pre-processing, custom conversion functions were applied to resolve encoding inconsistencies, and regular expressions were used to demarcate subject-predicate boundaries, followed by annotation using the “Begin (B)-Inside (I)-Other (O)” labeling scheme. The data were formatted into “word-part-of-speech-label” triplets to facilitate CRF model training. For feature engineering, we designed a word2features function to extract multi-dimensional features, encompassing word form, morphological variations, part-of-speech tags, contextual tokens within a window of  $\pm 2$  words and their combinations, sentence boundaries (BOS/EOS), and affixal information. The CRF model was implemented using the sklearn-crfsuite library, optimized with the L-BFGS algorithm, with regularization parameters set to  $C_1 = 0.2$  and  $C_2 = 0.1$ , and a maximum of 200 iterations. The dataset was partitioned into training and test sets in a 9:1 ratio, containing 35,635 and 3,960 samples, respectively. Experimental results indicate a weighted F1-score of 0.7459, with F1-scores of 0.7541 for I-labels and 0.6739 for B-labels. The weighted precision and recall were 0.7675 and 0.7257, respectively. While the model demonstrates strong performance in identifying internal elements of subject-predicate structures, there remains room for improvement in detecting initial boundaries and handling long-distance dependencies. This study confirms that integrating part-of-speech and composite contextual features enhances model performance, offering a valuable reference for the automated recognition of syntactic structures in Chinese.

## Keywords

Chinese Subject-Predicate Phrases, Conditional Random Fields (CRF), Automatic Identification, Feature Engineering

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

自然语言处理(NLP)作为人工智能的核心领域之一,旨在使计算机能够理解、解释和生成人类语言,从而实现人机之间更自然地交互。随着深度学习技术的迅猛发展,NLP在机器翻译、文本分类、情感分析等任务中取得了突破性进展。然而,在序列标注这一经典问题上,传统机器学习方法依然具有不可替代的价值。命名实体识别(NER)作为序列标注的典型应用,专注于从非结构化文本中自动识别和分类具有特定意义的实体,如人名、地名、机构名、时间表达等,是信息抽取、知识图谱构建、问答系统等高级应用的基础模块[1]。

近年来,基于深度学习的NER方法,特别是结合双向长短期记忆网络(BiLSTM)与CRF的混合架构,已成为工业界和学术界的主流选择。这类方法通过神经网络自动学习特征表示,避免了繁琐的人工特征工程。然而,这种优势也伴随着显著的成本:大规模标注数据需求、高昂的计算资源消耗、模型可解释

性差以及在小样本场景下的过拟合风险。相比之下,传统的机器学习方法,特别是条件随机场(CRF),凭借其在处理结构化数据时的独特优势,在特定应用场景下仍展现出强大的竞争力。CRF 通过显式特征工程融入人类语言学知识,模型轻量且易于部署,在数据规模有限、计算资源受限或需要强可解释性的领域,如古籍文献处理、低资源语言分析、教育辅助系统等,具有重要应用价值[2]。

汉语作为典型的分析型语言,缺乏形态变化,其语法关系主要依赖语序和虚词来表达,这给句法分析带来了独特的挑战。主谓结构作为汉语句子的基本语法骨架和核心命题载体,其准确识别对于深层语义理解至关重要。主谓结构不仅承载着句子的主体与动作关系,更是理解句子意图、抽取事件信息、构建语义角色框架的基础。然而,汉语主谓短语的自动识别面临多重困难:首先,汉语词界模糊,分词错误会级联影响句法分析;其次,主语和谓语成分长度不一,可以是单个词,也可以是复杂的短语结构,甚至存在主语省略和谓语残缺现象;再次,汉语中话题化、焦点移位等语用现象导致主谓成分可能不连续,形成长距离依赖;最后,兼语、连动、主谓谓语句等特殊句式增加了识别复杂性[3]。

针对上述挑战,研究聚焦于汉语主谓短语的自动识别问题,提出一种基于条件随机场(CRF)的解决方案。我们系统性地探索了特征工程在汉语句法分析中的有效性,旨在回答以下研究问题:(1) 在传统 CRF 框架下,何种特征组合对汉语主谓短语识别最为有效?(2) 词性信息与上下文特征如何协同提升识别性能?(3) 模型在处理边界识别和长距离依赖方面存在哪些局限?为回答这些问题,我们从清华-汉语句法树库(TCT/973 树库)中提取了 39,595 个标注样本构建高质量数据集。在模型构建上,我们设计了涵盖词法、句法、上下文等多维度的特征模板,并采用 L-BFGS 优化算法对 CRF 模型进行训练[4]。实验结果表明,该模型在主谓短语识别任务上取得了加权 F1 值 0.7459 的成绩,其中短语内部标记(I)的 F1 值达 0.7541,显著优于起始标记(B)的 0.6739。深入研究证实,词性特征与上下文信息的组合特征对性能提升贡献显著,但模型在处理长度超过 5 个词的长主谓结构时,F1 值下降约 12 个百分点,揭示了其在长距离依赖建模上的不足。本研究不仅验证了传统机器学习方法在特定汉语句法分析任务中的有效性,更为资源受限场景下的句法结构识别提供了可复用的方法论参考[5]。

## 2. 研究思路

文章采用“数据驱动 + 知识指导”的混合方法论,整体技术框架包含四个核心模块:语料获取与预处理、特征工程与表示、模型训练与优化、性能评估与错误分析。各模块环环相扣,形成完整的研究闭环[6][7]。

在数据层面,文章选择清华-汉语句法树库(TCT)作为训练语料。该语料库由清华大学自然语言处理实验室构建,基于 973 项目支持,包含约 100 万词的人工标注句法树,覆盖新闻、文学、学术、口语四大领域,标注体系融合成分结构分析与依存关系描述。我们提取其中主谓结构完整的句子,经清洗后构建平衡数据集,确保各领域样本分布均匀,避免领域偏差影响模型泛化能力。

特征层面,研究贯彻由浅入深、由局部到全局的设计原则。基础特征涵盖词形、词性等词汇级信息;上下文特征捕获句法环境的局部模式;组合特征建模词与词性、词与位置的交互;边界特征强化句子首尾的约束。特别地,引入词缀特征,利用汉语中后缀“者”“性”“化”等可指示词性功能的特点,增强模型对未知词的泛化能力,所有特征通过 one-hot 编码转换为二进制向量,构成 CRF 的观测序列输入。

模型层面,研究采用 sklearn-crfsuite 实现线性链 CRF。该库基于 CRFsuite 后端,支持 L1 和 L2 正则化,可有效防止过拟合,优化算法选择 L-BFGS,其在处理大规模稀疏特征时收敛速度快且内存效率高。通过开发集调参确定最优超参数组合,特别关注正则化系数的权衡:过高的正则化导致欠拟合,损失关键语言学信号;过低则易过拟合,在测试集上泛化性能下降。最终设置的  $C_1 = 0.2$  (L1)、 $C_2 = 0.1$  (L2)在偏差-方差间取得良好平衡。

最后，评估环节采用分层抽样划分训练集和测试集，确保主谓结构类型分布一致。评估指标除常规的精确率、召回率、F1 值外，还针对 B、I 标签分别计算性能，精准定位模型短板。在错误分析环节，随机采样 100 个错误案例进行人工标注，归纳错误类型并量化分布，为后续改进提供数据支撑，整体评估体系兼顾宏观性能与微观诊断，保证研究结论的可靠性与深度。

### 3. 数据预处理

#### 3.1. 数据库的选择

文章研究使用的数据来源于清华 - 汉语句法树库(TCT)。该语料库是一个大规模的中文句法树库，包含了丰富的句法结构标注信息[8]，特点如下：

该数据库包含大规模的标注数据，比如 TCT 语料库有数万条标注好的句子，涵盖多种句法结构，为本研究提供丰富的训练和测试数据；数据库有详细的句法标注，每条句子不仅标注了词性信息，还标注了主谓结构等句法关系，为文章的目标——识别主谓结构提供了强大的数据支持；其中包含多领域的文本，涵盖新闻、文学、口语等多个语体语域，具有较高的数据多样性[9]。

#### 3.2. 数据转换函数的设计

由于原始的 TCT 语料库数据格式复杂，且存在编码问题，因此需要进行格式化处理，以便用于 CRF 模型的训练。为此，研究使用数据转换函数 `change_copora` 对 TCT 的数据文本进行编码转换和修复。该函数作用如下：

修复编码问题：原始数据可能存在编码不一致的问题，导致读取时出现乱码。我们通过尝试多种编码方式，如 `gb18030`、`gbk` 等进行解码，并在解码失败时采用容错机制 `errors='replace'`，确保数据能够正确读取。

主谓结构的识别与标注：通过正则表达式匹配，识别出主谓结构的起始和结束标记，并将其标注为 B (开始)和 I (内部)标签，而对于非主谓结构的词语，则标注为 O (其他)。

数据格式化：将处理后的数据存储为适合 CRF 模型训练的三元组格式，即每行包含一个词语、词性及其对应的标签，句子之间用空行分隔。

### 4. 特征工程

在条件随机场(CRF)模型中，特征抓取是提升模型性能的重要步骤。特征抓取的目标是从原始数据中提取出对模型训练和预测有帮助的信息。对于中文主谓结构识别任务，合适的特征能够显著提高模型对主谓短语的识别能力。因此，研究设计出一套综合的特征提取方法，能够充分利用上下文信息和词性信息。

#### 特征提取函数

为了实现对中文主谓结构的识别，研究设计了一个特征提取函数 `word2features`，该函数针对每个词语提取以下特征：词本身及其变形(见表 1)、词性信息(见表 2)、上下文信息(见表 3)、句子边界信息(见表 4)以及词的前后缀(见表 5)。

**Table 1.** The functional feature table of a word and its inflections

**表 1.** 词本身及其变形的函数特征表

当前词的属性 Attributes of the Current Word	特征函数 Feature Function
当前词 Current Word	<code>word</code>
当前词的小写形式 Lowercase Form of the Current Word	<code>word. Lower ()</code>

续表

当前词的首字母是否大写 Whether the Initial Letter of the Current Word Is Capitalized	word. Istitle ()
当前词是否全为大写 Whether the Current Word Is All Uppercase	word. Isupper ()
当前词是否为数字 Whether the Current Word Is a Number	word. Isdigit ()

**Table 2.** Part-of-speech information feature table**表 2.** 词性信息特征表

词性特征 Part-of-Speech Feature	特征函数 Feature Function
当前词的词性 Part-of-Speech of the Current Word	postag
词性前缀 Part-of-Speech Prefix	postag [:2]

**Table 3.** Context information feature table**表 3.** 上下文信息特征表

上下文信息 Context Information	特征函数 Feature Function
前一个词及其词性 The Previous Word and Its Part-of-Speech	-1: word, -1: postag
后一个词及其词性 The Next Word and Its Part-of-Speech	+1: word, +1: postag
前两个词及其词性 The Previous Two Words and Their Part-of-Speech	-2: word, -2: postag
后两个词及其词性 The Next Two Words and Their Part-of-Speech	+2: word, +2: postag
相邻词的组合特征 Combination Feature of Adjacent Words	如-1 + 0: words, 0 + 1: words

**Table 4.** Sentence boundary information annotation table**表 4.** 句子边界信息标注表

句子边界 Sentence Boundary	特征函数 Feature Function
句子开始标记 Sentence Start Marker	BOS
句子结束标记 Sentence End Marker	EOS

**Table 5.** Feature table of a word's prefixes and suffixes**表 5.** 词的前后缀特征表

词的前后缀特征 Word's Prefixes and Suffixes Feature	特征函数 Feature Function
当前词的最后一个字符 The Last Character of the Current Word	word [-1]
当前词的最后两个字符 The Last Two Characters of the Current Word	word [-2:]
当前词的前三个字符 The First Three Characters of the Current Word	word [:3]

## 5. 模型训练与评估

### 5.1. 条件随机场模型概述

条件随机场(Conditional Random Fields, CRF)是一种判别式序列标注模型, 广泛应用于自然语言处理中的命名实体识别(NER)任务, 在引言部分提到, 这里不再赘述。CRF 模型通过定义条件概率分布来预测序列标签, 能够充分利用上下文信息和特征函数, 从而在序列标注任务中表现出色。其核心思想是通过特征函数将观察序列与标签序列联系起来, 最大化条件概率。



在本研究中，笔者使用了 `sklearn-crfsuite` 库来实现 CRF 模型。该库提供了高效的训练和预测功能，支持多种优化算法和正则化方法，适合处理大规模数据集。CRF 直接建模条件概率其概率形式为：

$$P(Y|X) = \frac{1}{Z(x)} \exp \left( \sum_{t=1}^T \sum_k \lambda_k f_k(y_t, y_{t-1}, x, t) \right)$$

其中  $P(Y|X)$ ， $X$  为观测序列， $Y$  为标签序列， $Z(x)$  为归一化因子， $f_k$  为特征函数， $\lambda_k$  为权重参数。

5.2. 划分训练集和测试集

为了训练 CRF 模型，笔者将数据集划分为训练集和测试集。  
数据集划分：为了科学评估模型的能力，笔者对经过预处理的完整数据集 39,595 个标注样本进行分层随机划分。采用 9:1 比例，即约 35635 个样本作为训练集，用于模型参数的学习与优化；剩余约 3960 个样本作为独立的测试集，严格保留至最终阶段，确保评估结果的客观性。

5.3. 特征提取

特征工程是 CRF 模型性能的关键，研究应用了在第四节特征工程中详述的特征提取函数 `word2features`。该函数作用于序列中的每一个词，根据预设的多层次特征，如词形、词性、上下文窗口、组合特征等，计算并生成对应的特征向量。最终，训练集和测试集中的每一个句子都被系统地转换为一个特征向量序列及其对应的标签序列，构成了模型可直接处理的输入格式[10]。特征向量的维度和内容是根据语言学的词汇、语法等知识设计的。

5.4. 模型参数设置

设置 CRF 模型的参数(见表 6)，包括训练算法、正则化系数、最大迭代次数等。

Table 6. Model parameter setting table  
表 6. 模型参数设置表

参数名称 Parameter Name	参数设置 Parameter Setting
算法 Algorithm	lbfgs
正则化系数 Regularization Coefficient	$C_1 = 0.2, C_2 = 0.1$
最大迭代次数 Maximum Iteration Count	200
允许所有可能的转移 Allow All Possible Transitions	<code>all_possible_transitions = True</code>
开启调试信息 Enable Debug Information	<code>Verbose = True</code>

其中，启用 `all_possible_transitions=True` 选项，允许模型在训练过程中学习任意状态标签，如 B、I、O 之间的转移概率，而不施加任何预定义的语法或标签序列规则限制。增强了模型的灵活性，使其能够自主地从数据中学习到的复杂的标签依赖关系。设置 `verbose = True` 选项，在训练过程中输出详细的迭代日志，包括每轮迭代的损失、梯度范数等信息，实时监控这些调试信息对于诊断训练过程是否正常以及辅助调整学习率等参数至关重要，保证训练过程的透明度和可控性。

5.5. 模型评估

研究采用条件随机场(CRF)模型进行汉语主谓结构识别任务，在清华大学 973 树库数据集上的实验结果如表 7 所示。模型整体性能达到加权 F1 值 0.7459，其中加权精确率为 0.7675，加权召回率为 0.7257，

表明模型在识别主谓结构时：精确率优于召回率，当模型预测某个词属于主谓结构时，模型判断具有较高可靠性，即 76.75%的正确率。

**Table 7.** Table of Chinese subject-verb structure recognition results

**表 7.** CRF 汉语主谓结构识别结果数据表

评价指标 Evaluation Metrics	加权平均值 Weighted Average	B 标签 B-Tag	I 标签 I-Tag
精确率 Precision	0.7675	0.7207	0.7728
召回率 Recall	0.7257	0.6328	0.7362
F1 值 F1-Score	0.7459	0.6739	0.7541

**Table 8.** CRF typical recognition errors of Chinese subject-verb phrases

**表 8.** CRF 汉语主谓短语典型识别错误

错误类型 Error Types	比例 Proportion	典型案例 Typical Cases
起始词误判 Misclassification of Starting Words	42.3%	重点(O/B)
边界识别错误 Boundary Recognition Errors	37.7%	(O/I)
副词短语误判 Misclassification of Adverbial Phrases	15.2%	也(O/I)
其他错误 Other Errors	4.8%	

识别完整性有待提升，模型对真实主谓结构的覆盖度不足，仅能识别 72.57%的实际主谓结构。样本不均衡问题显著，I 标签样本量 34676，是 B 标签样本量 3919 的 8.84 倍，导致模型对主谓结构起始位置的识别能力较弱。

如表 7 所示，模型对两类标签的识别存在显著差异，B 标签 F1=0.6739，I 标签 F1=0.7541。B 标签召回率仅为 63.28%，意味着 36.72%的主谓结构起始词未被正确识别，但 I 标签识别较优，F1 值达 0.7541，表明模型能较好识别主谓结构内部成分，但当起始词识别错误时，后续 I 标签预测会产生连锁错误。典型识别错误如表 8 所示。

## 6. 讨论与结论

### 6.1. 讨论

研究基于条件随机场(CRF)构建的汉语主谓短语识别模型，在测试集上取得加权 F1 值 0.7459、精确率 0.7675、召回率 0.7257 的性能，其中主谓结构内部成分(I 标签)F1 值达 0.7541，证实了词性与上下文组合特征在句法识别中的有效性。这一结果的核心价值在于，在无需大规模标注数据和高额计算资源的前提下，通过显式特征工程融入语言学知识，为资源受限场景提供了轻量、可解释的句法分析方案，其适用于古籍处理、低资源语言分析等深度学习模型难以落地的场景，凸显了传统机器学习方法在特定领域的不可替代性。但模型性能存在显著的结构性差异，I 标签识别效果优于 B 标签，且长距离依赖结构。这一现象并非单纯的参数优化问题，而是 CRF 模型本身所导致的，其作为线性链模型，本质上依赖局部特征窗口捕捉序列依赖关系，无法直接建模主谓成分不连续、跨距较大的语义关联，而汉语中话题化、焦点移位等语用现象恰好普遍存在此类长距离依赖，导致模型对非连续主谓结构的识别能力先天受限。

研究的特征工程带有完备性缺陷，word2features 函数虽涵盖词形、词性、上下文等多维度特征，但仍存在两点局限：一是特征设计依赖人工语言学知识，具有较强主观性，例如对词缀特征的选取仅覆盖

“者”“性”“化”等常见后缀，未纳入方言词缀、新兴网络词缀等特殊情况，导致对未知词的泛化能力不足；二是特征间的交互建模不足，当前采用简单组合特征(如-1+0:words)未能捕捉词性与上下文的高阶交互关系，而汉语主谓关系的判定往往依赖这种复杂交互，如副词修饰动词构成谓语时的词性——语义联动。

## 6.2. 结论

研究针对汉语主谓短语自动识别的挑战，提出基于条件随机场(CRF)的序列标注方法，通过构建含39595个标注样本的高质量数据集、设计涵盖词法、词性、上下文及词缀的特征工程框架，在测试集上取得加权 F1 值 0.7459 的性能，证实了词性与上下文组合特征在句法识别中的有效性，其核心贡献在于为数据规模有限、计算资源受限等场景，如古籍处理、低资源语言分析等提供了轻量、可复用的技术路径，所建立的“词语-词性-标签”三元组数据处理流程与特征模板也可为同类汉语句法识别任务提供参考，但需客观认知该方法的局限性——性能仍低于深度学习混合模型，且受 CRF 线性链架构约束，在长距离依赖、复杂句式及主谓结构起始边界识别上存在先天不足，样本不均衡与特征工程的主观性也进一步限制了模型泛化能力。未来研究可通过如“CRF+ 轻量神经网络”的混合架构突破线性建模局限[11]，结合无监督学习与句法依存特征实现特征工程的智能化升级，并深入探索模型在不同汉语句法结构中的适配机制，为传统机器学习方法在句法分析中的应用寻找优化方向。

## 基金项目

重庆市人文社科重点研究基地交通文化研究中心开放性课题资助，课题名称《中国传统交通文化术语知识图谱建设与传播研究》批准号 2024CTCR08。

## 参考文献

- [1] 桑德拉·库布利克, 舒巴姆·萨博, 李兆钧. GPT 使用 OpenAI API 构建 NLP 产品的终极指南[M]. 北京: 机械工业出版社, 2024.
- [2] 钱小飞. 汉语长名词短语识别研究[M]. 上海: 上海大学出版社, 2023.
- [3] 张金柱, 于文倩. 基于短语表示学习的主题识别及其表征词抽取方法研究[J]. 数据分析与知识发现, 2021, 5(2): 50-60.
- [4] 谢靖, 苏新宁, 沈思. CSSCI 语料中短语结构标注与自动识别[J]. 现代图书情报技术, 2012(12): 32-38.
- [5] 徐艳华. 基于语料库的基本名词短语研究[J]. 语言文字应用, 2008(1): 120-125.
- [6] 孔玲, 胡昊天, 张卫, 等. 跨学科知识扩散视域下学科交叉科学术语识别与特征计算[J]. 图书情报工作, 2024, 68(12): 119-137.
- [7] 陈禹, 刘林旭. 现代汉语主谓名素复现及其统计分析[J]. 语言文字应用, 2021(2): 79-88.
- [8] 姚从军, 罗丹. 面向信息处理的汉语主谓谓语句的组合范畴语法分析[J]. 中国社会科学院研究生院学报, 2019(2): 14-24.
- [9] 张岳, 滕志扬, 张梅山, 等. 自然语言处理基于机器学习视角[M]. 北京: 机械工业出版社, 2024.
- [10] 刘洪超, 黄居仁, 侯仁魁, 等. 基于语言学特征向量和词嵌入向量的汉语动词事件类型预测[J]. 中文信息学报, 2018, 32(1): 26-33.
- [11] 邱锡鹏. 神经网络与深度学习[M]. 北京: 机械工业出版社, 2024.