DOI：10.13718/j.cnki.xdzk.2024.10.019
张明，廖希. 基于人工智能神经网络的上下文介词消歧方法［J］. 西南大学学报(自然科学版)，2024，46(10)：222-232.

基于人工智能神经网络的上下文介词消歧方法

张明[1]，　廖希[2]

1. 成都职业技术学院 软件学院，成都 610041；2. 重庆邮电大学 通信与信息工程学院，重庆 400065

摘要：介词结构的分析难点在于如何对介词及其结构进行有效分类，挖掘其语义信息，并对介词结构进行有效的消歧处理. 为了应对这一难题，结合人工智能和神经网络技术，提出一种基于长短期记忆和注意力机制的树递归神经网络模型，旨在解决自然语言处理中的上下文介词消歧问题. 该模型通过引入注意力机制，将模型注意力集中在与介词含义相关的关键信息上. 首先，通过嵌入上下文解析树和上下文词向量，捕捉上下文词汇之间的语义关系. 然后，采用带有长短期记忆功能的树递归神经网络(Long Short-Term Memory Tree Recurrent Neural Network，Tree-LSTM)模型为树中的每个节点生成隐藏特征，并递归地跟踪树中不同分支上的传播来计算树节点的上下文表示. 最后，为了减少噪声对上下文中与介词含义相关的关键信息的影响，引入注意机制卷积神经网络(Attention-based Convolutional Neural Network，ACNN)，使模型专注于需要消除歧义的文档中的重要部分. 这种方式使模型能够自动选择并关注与当前介词含义最相关的词汇，从而提高消歧准确性. 实验结果表明：在 Semeval 2013 Task 12 词义消歧数据集上，该文提出的模型取得了 88.04% 的 *F*1-*score*，优于现有主流深度学习模型，验证了该文方法的有效性.

关　键　词：人工智能；神经网络；介词消歧；深度学习；注意力机制

中图分类号：TP393　　**文献标志码**：A　　开放科学(资源服务)标识码(OSID)：

文 章 编 号：1673-9868(2024)10-0222-11

A Context Preposition Disambiguation Method Based on Artificial Intelligence Neural Network

ZHANG Ming[1]，　LIAO Xi[2]

1. *School of Software，Chengdu Polytechnic，Chengdu 610041，China*；
2. *School of Communication and Information Engineering，Chongqing University of Posts and Telecommunication，Chongqing 400065，China*

Abstract：The analysis of prepositional structures presents challenges in effectively classifying prepositions

收稿日期：2023-12-18
基金项目：国家自然科学基金项目(61801062)；重庆市自然科学基金项目(cstc2021Jcyj-msxmX0634).
作者简介：张明，硕士，副教授，主要从事人工智能应用研究.

and their structures, mining their semantic information, and effectively disambiguating prepositional structures. To address this challenge, this paper proposes an ACNN-Tree-LSTM model that combines artificial intelligence and neural network techniques, aiming to solve the problem of context-based preposition disambiguation in natural language processing. The core idea is to introduce an attention mechanism to focus the model's attention on key information in the context, which is relevant to the meaning of the preposition. In this study, the context parsing tree and context word embeddings were first embedded to capture the semantic relationships between context words. Then, the Tree-LSTM model was utilized to generate hidden features for each node in the tree, and the context representation of tree nodes was computed by recursively tracking propagation along different branches of the tree. Finally, to reduce the influence of noise on key information related to the meaning of the preposition in the context, an attention mechanism was introduced to enable the model to focus on the crucial parts of the reference document that require disambiguation. This approach allows the model to automatically select and pay attention to the vocabulary mostly relevant to the current prepositional meaning, thereby improving disambiguation accuracy. Experimental results on the Semeval 2013 Task 12 Word Sense Disambiguation dataset demonstrated that the proposed model achieved an F1-score of 88.04%, outperforming existing mainstream deep learning models and validating the effectiveness of the proposed approach.

Key words: artificial intelligence; neural networks; preposition disambiguation; deep learning; attention mechanism

随着人工智能领域的迅猛发展，自然语言处理(Natural Language Processing，NLP)作为一个重要的研究方向受到了广泛关注. NLP 旨在使计算机能够理解和处理人类语言，其中上下文理解是其中一个重要的任务. 在自然语言中，词义消歧(Word Sense Disambiguation，WSD)逐步引起了研究者们的关注[1].

由于神经网络强大的非线性拟合能力，因此深度学习模型被广泛用于词义消歧. 例如，何春辉等[2]基于余弦相似度，对未标记的语料使用长短期记忆(Long Short-Term Memory，LSTM)模型结合有标注语料上下文语境进行消歧. Arshey 等[3]选用多义词邻接的 4 个词的词形、词性和语义作为特征，使用深度信念网络(Deep Belief Networks，DBN)模型来消歧. Chauhan 等[4]将深度学习模型应用于词义消歧，直接优化了给定相似度的文档，在无监督预训练阶段使用叠加去噪自动编码器来学习初始文档表示，最后进行微调. 由于短文本语义的稀疏性，需要深度神经网络来进一步探索语义，但过度的深度堆叠自动编码器容易出现梯度消失问题. Loureiro 等[5]利用堆叠双向长短期记忆网络(Bidirectional Long Short-Term Memory，Bi-LSTM)的双重关注机制，从上下文特征、词义描述特征等方面计算词义之间的关联性. 然而，该模型侧重于句子的长距离依赖，没有考虑文本的局部特征. Li 等[6]根据双向 LSTM 编码器捕获了提及的词汇、句法和本地文本信息，并使用卷积神经网络(Convolutional Neural Network，CNN)与细粒度类型的结构化信息源相结合对实体文档进行建模. 结构化信息包括对知识库的描述. 知识库的质量直接影响到实体消歧结果，而该研究没有考虑到数据集中实体与知识库中的实体不一致的情况. 段宗涛等[7]提出了用于词义消歧的成对连接. 通过模拟 Kruskal 算法，使用配对连接算法来近似解决 MINTREE(基于树的词义消歧目标)问题，并使用 Word2vec 中的跳格方法来完成文本矢量化表示. 这种方法生成的词向量与词本身一一对应，不能反映同一词在不同语境中的真实含义. Alokaili 等[8]使用一个联合排名框架来寻找相似或相关的实体以消除歧义，他们提出用一个词义消歧框架来扩展概念性的短文嵌入模型，并使用注意力模型来选择相关的词进行预测. 然而，在处理短文 NLP 任务时文本中包含的有用信息相对较少，仅靠注意力机制无法获得完整的语义知识.

介词属于助词，在语言中所占的比例并不大，但却是一个重要而常见的词类. 它是虚词中常见的一类，用来表示词与词或句之间的关系，介词不能单独作句子成分，需要与其他实词共同构成介词短语，作用是在句子中修饰、补充谓语，揭示与动作、性状等有关的如时间、地点、比较、施事、受事、对象、方式等.

上下文介词消歧的重要性在于，它在多个NLP任务中扮演着关键角色，包括句法分析、语义理解、机器翻译和信息检索等，准确地识别上下文介词的含义对于这些任务的性能和效果具有重要意义. 然而，由于上下文的复杂性和多义性，上下文介词消歧仍然是一个具有挑战性的问题. 当前的研究主要集中在传统的基于规则和统计的方法上，这些方法通常依赖于手工特征工程和人工定义的规则，限制了其在复杂上下文中的适应能力. 此外，这些方法往往无法充分利用大规模数据和深度学习的优势，因此在处理复杂语义场景时存在一定的局限性.

为了解决上述问题，本文结合人工智能和神经网络技术，提出一种带有长短期记忆基于注意力机制的树递归神经网络模型(ACNN-Tree-LSTM)用于上下文介词消歧，旨在通过深度学习模型的应用，提高消歧任务的准确性和泛化能力. 实验验证了本文模型能够更好地利用上下文信息，并有效地解决了介词歧义问题. 通过与其他先进的深度学习模型在大型词义消歧数据集上的比较结果表明，本文模型在上下文介词消歧任务中具有非常明显的优越性.

本文的研究具有以下几个方面的意义：

方法创新：本文探索并提出了一种新颖的基于人工智能神经网络的上下文介词消歧方法，使用Tree-LSTM作为文本表示，通过树递归神经网络捕获语义信息，并添加自注意力机制来进一步学习特征信息. 该方法将深度学习模型与上下文建模相结合，在介词消歧领域提供了一种新的解决方案，并为上下文理解任务的研究和应用提供了新的思路.

提高任务准确性：本文的方法旨在提高上下文介词消歧任务的准确性. 通过充分利用深度学习模型的学习能力和泛化能力，尽可能减少消歧任务中的误判和错误分类，从而提高整体任务的准确性.

1 文献综述

1.1 介词概念及其歧义分析

在英语中，介词可分为简单介词、复合介词和短语介词等多种类型，在语义分析中发挥着重要作用[9-10]，例如：at、by、in、of、for、off、from是简单的介词. 复合介词是由一个介词和一个副词或形容词组成的固定搭配，它们的含义是整体构成短语所表示的含义，例如：according to，because of，in front of，from under等. 短语介词是由多个词组成的固定短语，在语法上作为一个介词来使用. 短语介词可以是介词、名词或代词的组合，例如：at first，on time，in school等.

介词通常具有多个含义，其具体含义取决于上下文的语境. 例如，在句子“I ran____the park”中，介词“through”和“to”的含义完全不同，分别表示“穿过”和“到达”的意思.

上下文介词消歧是指在给定句子中确定介词具体含义的任务. 例如：

“Buy a car with a steering wheel”.

句中的介词短语“with a steering whee”修饰了名词“a car”. 在该句中，正确的理解是人们应该买一辆带方向盘的汽车，而不是说人们应该用方向盘作为购买汽车的交易条件.

上下文介词歧义问题也可以指一个特定的介词，在不同语义场景下具有多个不同含义. 如图1，图2两种情况.

在图1句子1中，介词短语“with a fork”修饰在动词“eats”上，表示用于进食动作的工具(叉子). 图2中的句子2看起来似乎与第一个句子只有很小的区别，但是从图1和图2中的语法层次树可以看出，它们的语法结构有很大不同. 介词短语“with apple”不修饰在动词上，而是修饰在名词“pizza”上，表示和披萨一

起进食的食物(苹果).

因此，准确理解上下文中介词的含义对于正确解析句子的语义至关重要.

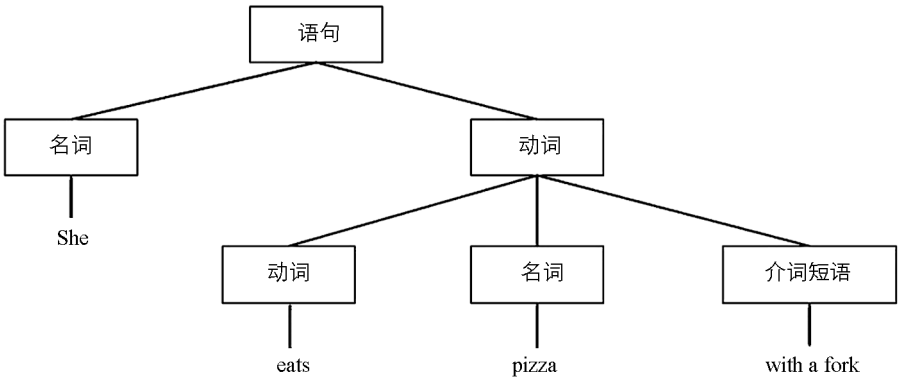

图 1　句子 1

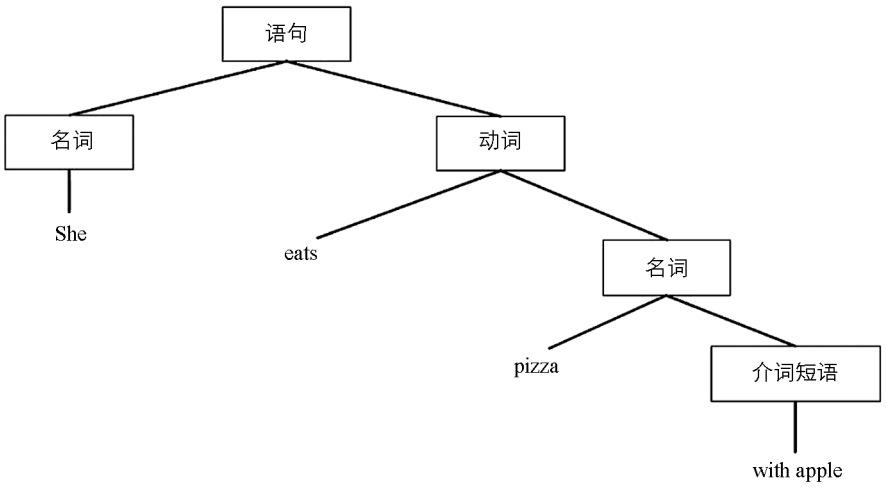

图 2　句子 2

1.2　介词消歧研究现状

1.2.1　基于规则的介词消歧方法

早期的研究主要采用基于规则的方法来解决介词消歧问题. 基于规则的方法依赖于手工定义的规则和语法知识，通过匹配和推理来确定介词的含义. 这些方法在小规模数据集上表现良好，但在复杂语义场景下的适应能力有限.

一种常见的基于规则的方法是使用词典或语法规则来确定介词含义. 研究者们根据词典中的定义和示例，或者基于句法规则，手动为每个介词定义含义，然后根据上下文进行匹配. 然而，这种方法需要大量的人工工作和专家知识，并且难以泛化到新的语义环境中.

1.2.2　基于统计的介词消歧方法

统计方法通过统计语言模型和机器学习算法来解决介词消歧问题. 其中，最常用的方法是基于特征工程的机器学习方法，通过选择和转化特征来训练分类器进行消歧[11]. 这些方法通常依赖于手工定义的特征和特征选择算法，限制了其在复杂语义场景下的准确性和泛化能力.

特征工程的方法通常会考虑上下文中的词汇、句法结构和语义信息等特征. 例如，可以考虑上下文中的词性、词频、上下文窗口中的其他词汇等特征. 然后，通过选择合适的特征和使用机器学习算法，如支持向量机(Support Vector Machine，SVM)或朴素贝叶斯分类器来进行消歧.

1.2.3　基于深度学习的介词消歧方法

近年来，随着人工智能、神经网络技术的兴起，基于深度学习的方法在介词消歧问题中取得了显著的

进展. 这些方法主要通过将上下文建模和语义信息捕捉纳入神经网络的训练过程中，从而实现了对上下文介词含义的准确判断. 深度学习模型，如神经网络和 Transformer 等具有强大的表征能力和学习能力，能够从大规模数据中学习复杂的语义信息.

一种常用的方法是使用预训练语言模型，如来自 Transformers 的双向编码器表示(Bidirectional Encoder Representations from Transformers，BERT)，通过无监督学习方式学习词向量和语言模型，然后在消歧任务上进行微调. BERT 模型能够通过双向上下文信息的学习，捕捉到词语的丰富语义表示[12-13]. 研究者们在 BERT 模型的基础上进行了改进和优化，如使用不同的注意力机制和网络结构，以提高模型在上下文词义消歧任务上的性能[14].

除了 BERT，还有其他基于深度学习的模型被应用于介词消歧任务，如卷积神经网络(CNN)、循环神经网络(Recurrent Neural Network，RNN)、长短期记忆网络(LSTM)等. 这些模型通过多层神经网络结构，能够捕捉上下文的局部依赖关系和长期依赖关系，从而更好地理解上下文中介词的语义含义.

此外，一些研究工作还尝试将注意力机制引入上下文介词消歧任务中. 注意力机制能够自动地将注意力集中在关键信息上，有助于模型更好地理解上下文的重要部分. 通过引入注意力机制，研究者们能够更准确地捕捉到上下文中与介词含义相关的信息，提高了模型的消歧性能.

2 本文方法

2.1 上下文表示

树递归神经网络(Tree-RNN，TNN)[15]被引入利用文本的语言结构来组成句子表示. 通过计算两个候选子节点的表示(如果两个节点合并)，网络将计算其父节点的表示以及新节点的合理性分数. 然后递归地重复这个过程，可以获得短语的语义表示. 本文选择带有长短期记忆功能的树递归神经网络(Long Short-Term Memory Tree Recurrent Neural Network，Tree-LSTM)来编码上下文信息，以便通过考虑小元素的语义组成来获得更好的文本单元表示. 与 TNN 连接时，LSTM 还可随着时间的推移保留序列信息. TNN 可以被视为 RNN 的超集，具有线性结构. RNN 需要先前的上下文来捕获短语，并且通常会捕获最终向量中过多的最后一个单词. 与 RNN 不同，TNN 则可以从小块中捕获更大的类似于语言结构的结构.

Tree-LSTM 模型的目标是通过递归地跟踪树中不同分支上的传播来加强树节点的高级表示. 在 Tree-LSTM 结构中，如果一个单词的颜色比另一个单词深，则意味着该单词与上下文向量的相关性比其他单词更大. 如图 3 所示，句子“I moved into my flat on the 21st of last month.（我上个月 21 日搬进了我的公寓)”. 在这个树状结构中，每个单词都表示为一个节点，连接节点的线表示它们之间的语法关系. 重要介词“into” “on” “of”在最终的上下文向量中得到加强.

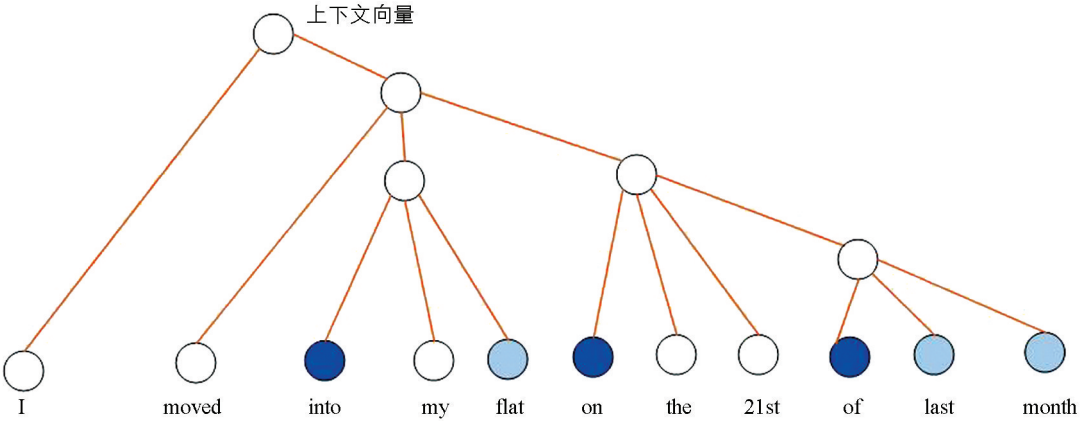

图 3 Tree-LSTM 结构句例

在上下文表示中，我们嵌入上下文解析树和上下文词向量. 首先使用 GloVe 向量作为词向量，然后使用 Stanford Parser 语法分析器来编码解析树. 嵌入后，使用选区 Tree-LSTM 为树中的每个节点生成隐藏特征. 在这个过程中，每个节点都会递归地处理其子节点的信息，然后更新自己的隐藏状态和存储单元，从而构建出反映整个句子语义的上下文向量.

2.2　文档和描述表示

在生成选区 Tree-LSTM 之后，根据子节点的表示递归地计算非叶节点的上下文表示，如图 4 所示.

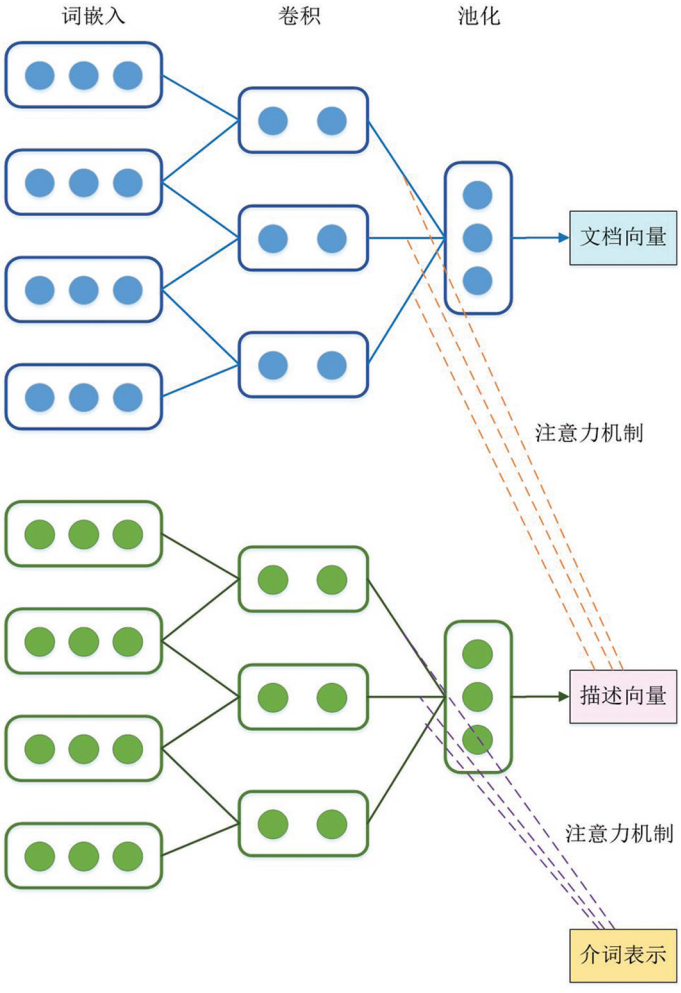

图 4　文档和描述表示概述

为了体现文档信息和实体描述页面，本文使用带有注意力机制的卷积神经网络(CNN)来编码这些背景信息. CNN 在处理数据时能够有效地提取出具有稳定性和抽象性的特征. 词嵌入之后，使用 CNN 为文档生成固定长度的向量. 本文使用修正线性单元(ReLU)作为激活单元，并将结果与最大池化相结合. 然而，CNN 无法捕获文档中的重要组成部分. 为了减少噪声对相关信息的影响，引入了一种注意机制，以使模型专注于需要消除歧义的相关信息的重要部分.

对于候选描述部分，使用介词表示作为卷积层输出的注意力权重来强化描述页面的关键部分. 介词表示通过候选介词的词向量的平均值来计算. 对于文档部分，使用候选介词描述向量作为注意力权重来加强文档中的重要组成部分.

2.3 类型表示

细粒度类型为介词提供结构化信息. 细粒度类型是指在分类任务中对目标进行更细致划分和分类的一种方法，它在分类任务中引入了更多的子类或更细致的分类标签，使得分类更具有精确性和详细性. 例如，对于动物分类任务，可以将目标进一步细分为猫科动物、犬科动物、鸟类等，这样的细粒度分类可以更准确地捕捉目标的特征和属性，并提供更丰富的语义信息. 通过引入更多的细粒度类型，可以提高分类模型的准确性和区分能力，使其能够区分具有相似特征的目标. 为了键入候选介词，本文使用自然语言处理工具包(Natural Language Toolkit，NLTK)，这是一个常用的Python库，提供了丰富的文本处理功能，包括句法分析和词性标注等，为介词消歧任务提供支持. NLTK为每个候选介词返回一组类型，然后计算每种类型与介词相关的概率.

本文需要上下文信息来确定每个细粒度类型. 通过使用一种特定的系统，利用一种专注于细节的编码神经模型来预测细粒度类型.

2.4 介词消歧

在执行介词消歧(Preposition Disambiguation，PD)系统之前，本文需要生成一些候选介词，这些候选介词具有先前得分，它们在之前的评估中已经得到了分数，而这些分数是通过使用提前计算好的频率字典得出的. 本文从多个细粒度计算候选介词之间的语义相似性. 公式(1)的特征 $F(w, e)$ 表示语义相似度的多个粒度.

$$\begin{aligned} F(w, e) = [&\cos(w_c, e_c), \cos(w_c, e_d), \cos(w_c, e_n), \cos(w_d, e_c), \\ &\cos(w_d, e_d), \cos(w_d, e_n), \cos(w_n, e_c), \cos(w_n, e_d), \cos(w_n, e_n)] \end{aligned} \tag{1}$$

公式(1)计算了两个词向量 w 和 e 之间不同组成部分的余弦相似度. 具体来说，w 和 e 被分解为它们的子组成部分，分别标记为 w_c, w_d, w_n 和 e_c, e_d, e_n. 本文PD系统的最终分数是先前分数和语义相似度分数的组合. 然后，选择最终得分最高的候选介词作为消歧系统的结果.

2.5 模型训练

由于TNN利用语言信息来表示句子，因此本文选择BRNN-CNN系统. 由于PD模型使用深度神经网络，因此通过PD模型的损失函数并最小化结果函数来进行训练.

$$\Theta' = \underset{\Theta}{\operatorname{argmin}} L_{PD} \tag{2}$$

其中，Θ 是PD任务的参数集，L_{PD} 为损失函数. 在这个过程中，目标是确定模型参数 Θ 的最优值，使损失函数 L_{PD} 达到最小值. 优化后的模型参数为 Θ'，它代表了模型训练完成后的参数状态.

2.6 模型构架概述

本文使用Tree-LSTM对上下文部分进行建模，因为Tree-LSTM在表示长句子的语义方面表现良好. 文档部分使用带有注意力机制(ACNN)的CNN进行编码. 为了消除噪声对文档部分的影响，采用注意力机制来捕获输入的重要组成部分. 对于类型部分，使用细粒度介词类型系统为每个介词提供一系列类型，以减少歧义.

在上下文介词消歧任务中，本文计算提及部分(上下文、文档、类型)和候选介词部分(上下文、文档、类型)之间的语义相似性，如图5所示.

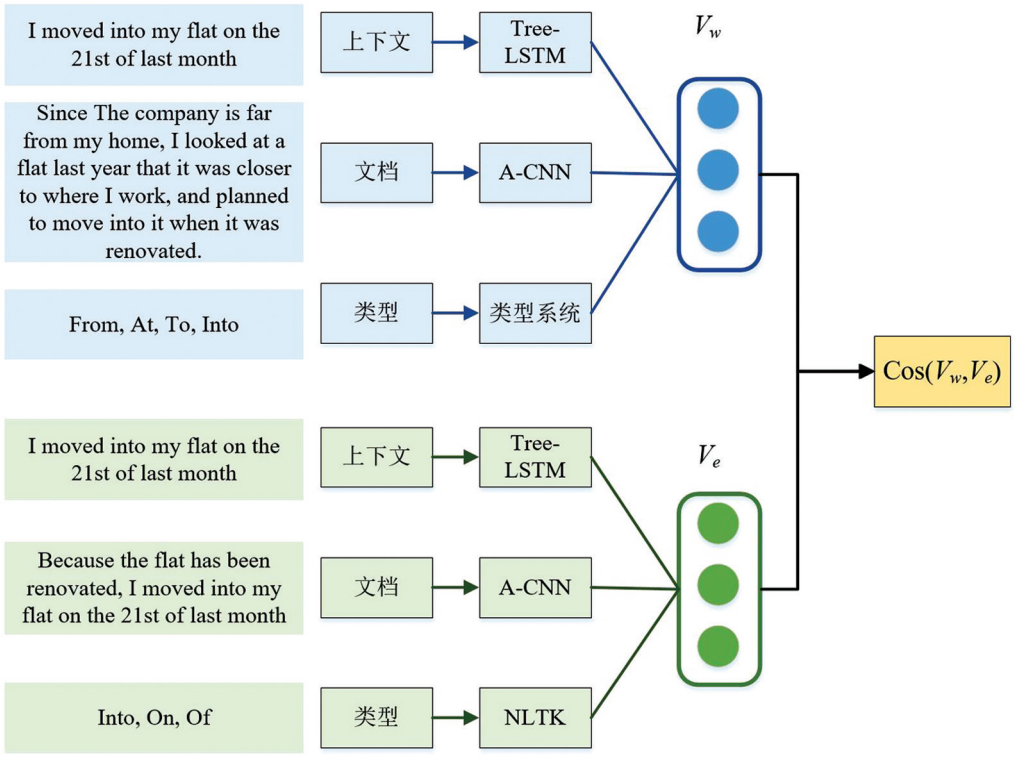

图 5　本文提及部分和候选部分介词消歧方法概述

3　实验设计与结果

3.1　数据集

在介词消歧领域，目前还没有专门针对介词消歧而构建的独立数据集. 因此，本文选择了一个常用的词义消歧数据集 Semeval 2013 Task 12. Semeval 2013 Task 12(数据集是从英文语料库中提取的). 该数据集包含 10 000 个标记的词义消歧实例，其中 5 134 个用于训练，4 866 个用于测试. Semeval 2013 Task 12 数据集覆盖了多个语义类别，包括动词、名词、形容词和介词.

3.2　评价指标

为了衡量模型在介词消歧任务中的性能，本文使用精度(Precision，P)、召回率(Recall，R)和 $F1$-*Score* 作为评价指标. 具体计算如式(3)、式(4)、式(5)所示.

$$P=\frac{TP}{TP+FP} \tag{3}$$

$$R=\frac{TP}{TP+FN} \tag{4}$$

$$F1=\frac{2\times P\times R}{P+R} \tag{5}$$

其中，TP 为被模型预测为正类的正样本；TN 为被模型预测为负类的负样本；FP 为被模型预测为正类的负样本；FN 为被模型预测为负类的正样本.

3.3　实验结果分析

本文模型在上下文介词消歧任务中与其他先进的深度学习模型(Word2vec-BiLSTM[16]、XLNet BERT[17]、Pre-trained BERT[18])进行了比较，在 3 个标准指标上的对比结果如图 6 所示.

图 6 中，对于 Word2vec-BiLSTM，其精度 P 为 75.76%，召回率 R 为 75.63%，$F1$ 值为 78.82%，位置特征的加入可以帮助模型感知文本中语义向量的位置信息，更好地理解数据中的语义关系. 与

Word2vec-BiLSTM相比，Pre-trained BERT中词之间的关系在P、R、$F1$方面分别提高了3.47%、3.74%和1.26%. 因为BERT生成的词向量是动态的，可以对单词的多义现象进行建模，能够更准确地反映词在当前语义中的实际含义. XLNet BERT通过常用的均值池化和最大池化融合特征来表示文本. 最大池化提取文本中最重要的特征；均值池考虑邻域中的所有语义信息. 该模型的性能相比Pre-trained BERT有很大的提高，分别高出6.80%、6.47%、5.35%. 而本文模型的P、R、$F1$值分别为89.25%、88.76%、89.58%，高于所有对比模型. 这是因为本文模型使用Tree-LSTM作为文本表示，通过树递归神经网络捕获语义信息，并添加自注意力机制来进一步学习特征信息，而不是简单地扁平化，通过自注意力为不同的语义分配权重，以更好地获得语义依赖关系. 此外，本文模型使用不同大小[3，4，5]的卷积核来提取BERT输出的文本语义信息，采用修正线性单元(ReLU)作为激活单元，并将结果与最大池化相结合来筛选提取的语义信息. 通过这两种方式的结合，可以有效地融合最重要的特征和邻域中的所有语义信息. 由图6可知，本文提出的具有注意力机制的树递归神经网络，在上下文介词消歧任务中取得了比现有主流模型更好的实验结果，验证了模型的有效性.

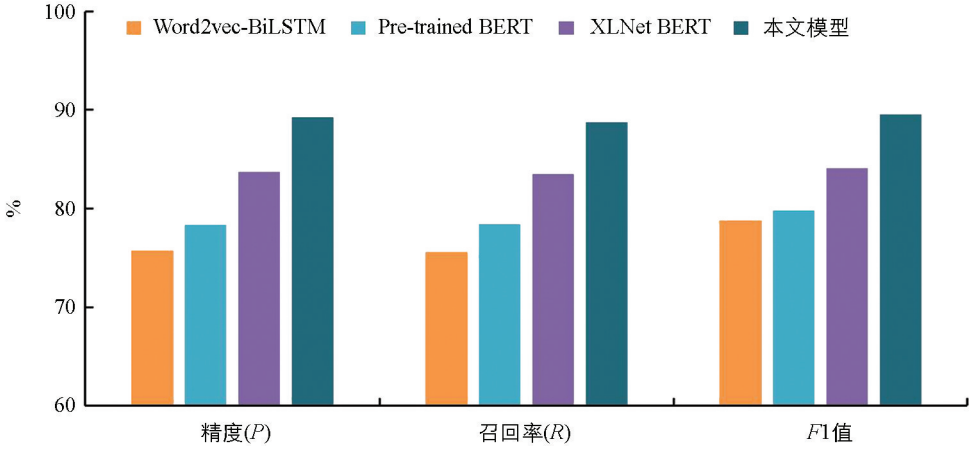

图6 不同模型在3个标准指标上的结果

为了进一步研究本文提出的Tree-LSTM模型中的神经网络与深度学习模型对上下文介词消歧的作用，分别使用学习率、树节点数、卷积核大小进行消融实验分析.

3.3.1 学习率对实验性能的影响

学习率(Learning rate)代表了神经网络中随时间推移和信息累积的速度. 学习率作为监督学习以及深度学习中重要的超参数，决定着目标函数能否收敛到局部最小值以及何时收敛到最小值. 学习率是最影响性能的超参数之一，合适的学习率能够使目标函数在合适的时间内收敛到局部最小值. 为了探究不同学习率对模型性能的影响，本文分别采用0.001、0.0001、0.00005的学习率进行实验，结果如表1所示.

表1 比较不同学习率的效果

学习率	精度(P)/%	召回率(R)/%	$F1$值/%
0.001	84.50	83.81	84.00
0.000 1	88.14	87.98	88.04
0.000 05	85.41	84.06	84.32

由表1可知，当学习率为0.000 1时，本文模型效果最好，P、R、$F1$值分别为88.14%、87.98%、88.04%. 然而，当学习率增加到0.001时，模型的结果P、R、$F1$分别下降了3.64%、4.17%、4.04%，因为当学习率过大时，模型会跨越最优值导致长时间不收敛，训练效果不佳. 当学习率为0.000 05时，模型的结果在P、R、$F1$方面分别下降了2.73%、3.92%、3.72%，因为当学习率太小时，模型容易下降陷入局部最优，无法达到全局最优点. 由此可知，本文模型的最佳学习率为0.000 1.

3.3.2　树节点数量对实验性能的影响

为了探讨递归神经网络中树节点数量对实验结果的影响，本文将树节点数量设置为 2、3、4，实验结果如表 2 所示.

表 2　树节点数量效果比较

树节点	精度(P)/%	召回率(R)/%	$F1$ 值/%
2	87.25	85.30	85.59
3	88.14	87.98	88.04
4	86.62	86.60	86.61

如表 2 所示，当使用 3 个树节点时，模型表现最佳. 当树节点数量设置为 2 时，模型 P、R、$F1$ 分别为 87.25%、85.30%、85.59%，相应指标分别下降 0.89%、2.68%、2.45%，这是因为树节点的数量太少，模型将无法充分拟合数据集. 当树节点数量为 4 时，模型 P、R、$F1$ 分别为 86.62%、86.60%、86.61%，相应指标分别下降 1.52%、1.38%、1.43%. 这是因为过多的树节点提取语义信息，导致模型过拟合，降低模型的迁移能力.

3.3.3　不同卷积核大小对实验性能的影响

为了探讨卷积核大小对模型性能的影响，本文选择大小为[2，3，4]、[3，4，5]和[4，5，6]的卷积核，实验结果如表 3 所示.

表 3　比较不同卷积核大小的效果

卷积核大小	精度(P)/%	召回率(R)/%	$F1$ 值/%
[2，3，4]	86.94	86.36	86.51
[3，4，5]	88.14	87.98	88.04
[4，5，6]	86.57	86.04	86.19

如表 3 所示，当卷积核大小为[3，4，5]时，模型达到最佳效果，P、R、$F1$ 分别为 88.14%、87.98%、88.04%. 当卷积核大小为[2，3，4]时，窗口较小，使得感受野的范围较小，感受野包含的文本信息较少，特征也有限，因此模型无法完全表达句子的这些特征，导致 P、R、$F1$ 值下降 1.20%、1.62%、1.53%. 当卷积核大小为[4，5，6]时，由于感受野的变化使得全局特征明显，模型获得的文本语义信息中存在相当大的噪声，模型复杂度增加. 模型的 P、R、$F1$ 分别下降了 1.57%、1.94%、1.85%.

4　结论

近年来的研究发现，歧义消除和理解主要集中在实词(代词、名词、动词等)的消解，对于高频虚词如介词的歧义消除和理解研究较为有限. 因此，本文提出一种基于人工智能的上下文介词消歧方法，采用带有长短期记忆功能的树递归神经网络模型. 该方法使用带有 LSTM 的 TNN(Tree-LSTM)对上下文进行建模，并通过注意力机制捕捉关键信息，以减少噪声对介词含义的影响. 通过多层神经网络结构，模型能够捕捉上下文中的局部依赖关系和长期依赖关系，更好地理解介词的语义含义. 引入注意力机制能够更准确地捕捉与介词含义相关的信息，从而提高消歧性能. 实验结果证明，本文模型在准确性和泛化能力方面具有优越性. 本文为相关领域的研究人员提供了有益的参考和启发，推动了上下文介词消歧技术的进一步发展和应用. 此外，本文模型还表现出较强的鲁棒性，能够适应不同领域和语义场景下的消歧需求. 然而，我们也注意到该模型在处理长文本和复杂语义场景时仍存在一定的挑战，这可能是注意力机制的局限性所致. 未来的研究将进一步探索改进注意力机制，以提升模型在复杂语境下的性能.

参考文献:

[1] 淑娴陈. 词义消歧研究综述[J]. 教育科学发展, 2022, 4(1): 137-139.

[2] 何春辉, 胡升泽, 张翀, 等. 融合深层语义和显式特征的中文句子对相似性判别方法[J]. 中文信息学报, 2022, 36(9): 28-37.

[3] ARSHEYM, ANGEL VIJI K S. An Optimization-Based Deep Belief Network for the Detection of Phishing E-Mails [J]. Data Technologies and Applications, 2020, 54(4): 529-549.

[4] CHAUHAN S, DANIEL P, SAXENA S, et al. Fully Unsupervised Machine Translation Using Context-Aware Word Translation and DenoisingAutoencoder [J]. Applied Artificial Intelligence, 2022, 36(1): 1771-1795.

[5] LOUREIRO D, REZAEE K, PILEHVAR M T, et al. Analysis and Evaluation of Language Models for Word Sense Disambiguation [J]. Computational Linguistics, 2021, 47(2): 387-443.

[6] LI J, SUNA X, HAN J L, et al. A Survey on Deep Learning for Named Entity Recognition [J]. IEEE Transactions on Knowledge and Data Engineering, 2022, 34(1): 50-70.

[7] 段宗涛, 李菲, 陈柘. 实体消歧综述[J]. 控制与决策, 2021, 36(5): 1025-1039.

[8] ALOKAILI A, ELBACHIRMENAIM. SVM Ensembles for Named Entity Disambiguation [J]. Computing, 2020, 102(4): 1051-1076.

[9] MADE JULIARTA I. Prepositional Phrase and Its Translations Found in the Novel "Budha, a Story of Enlightenment" [J]. E-Journal of Linguistics, 2021, 5(1): 28-47.

[10] NGHI T T, THANG N T, PHUC T H. An Investigation into Factors Affecting the Use of English Prepositions by Vietnamese Learners of English [J]. International Journal of Higher Education, 2020, 10(1): 24-40.

[11] 王奥, 吴华瑞, 朱华吉. 基于特征增强的多方位农业问句语义匹配[J]. 西南大学学报(自然科学版), 2023, 45(6): 201-210.

[12] 范齐楠, 孔存良, 杨麟儿, 等. 基于BERT与柱搜索的中文释义生成[J]. 中文信息学报, 2021, 35(11): 80-90.

[13] 陆伟, 李鹏程, 张国标, 等. 学术文本词汇功能识别——基于BERT向量化表示的关键词自动分类研究[J]. 情报学报, 2020, 39(12): 1320-1329.

[14] 张国标, 李鹏程, 陆伟, 等. 多特征融合的关键词语义功能识别研究[J]. 图书情报工作, 2021, 65(9): 89-96.

[15] 宋晓涛, 孙海龙. 基于神经网络的自动源代码摘要技术综述[J]. 软件学报, 2022, 33(1): 55-77.

[16] YUE W, LI L. Sentiment Analysis Using Word2Vec-CNN-BiLSTM Classification [C] //2020 Seventh International Conference on Social Networks Analysis, Management and Security (SNAMS). Paris: IEEE, 2020.

[17] RADKE M A, GUPTA A, STOCK K, et al. Disambiguating Spatial Prepositions: The Case of Geo-Spatial Sense Detection [J]. Transactions in GIS, 2022, 26(6): 2621-2650.

[18] PAWAR S, THOMBRE S, MITTAL A, et al. Tapping BERT for Preposition Sense Disambiguation [EB/OL]. 2021: arXiv: 2111. 13972. http://arxiv.org/abs/2111.13972.

责任编辑 夏娟