

计算机应用中人工智能 自然语言处理技术的优化与实践

文 ◆ 广州达到信息技术有限公司 蔡丽芸

引言

随着信息技术的迅猛发展，人工智能（AI）已成为推动社会进步和技术创新的核心力量。自然语言处理（NLP）作为人工智能领域的重要分支，致力于让计算机理解和生成人类的自然语言，进而实现人机之间的自然交互。近年来，自然语言处理技术在计算机应用中得到广泛应用，涵盖智能客服、机器翻译、文本分类、情感分析等诸多领域，极大地提升了信息处理的效率和质量^[1]。然而，在实际应用过程中，自然语言处理技术仍面临诸多挑战。自然语言本身具有复杂性、多样性和模糊性，使得机器难以完全准确地理解和生成自然语言。此外，数据质量参差不齐、模型性能有待提高以及计算效率较低等问题，也制约着自然语言处理技术的进一步发展。因此，对自然语言处理技术进行优化研究，不仅具有重要的理论意义，还有着迫切的现实需求。本文将从自然语言处理技术的基本原理出发，深入探讨其在计算机应用中的优化策略，并通过实际案例分析优化后的技术应用效果，以期对相关领域的研究和实践提供有益参考。

1 自然语言处理技术概述

1.1 自然语言处理的定义与任务

自然语言处理（NLP）是计算机科学、人工智能和语言学等多学科交叉融合的领域，其核心目标是使计算机具备理解、生成和处理人类自然语言的能力^[2]。该领域的主要任务丰富多样，文本分类旨在将文本划分到预定义的类别中，如新闻分类、垃圾邮件检测；情感分析用于判断文本所表达的情感倾向，包括正面、负面或中性；机器翻译实现不同语言文本之间的自动转换；问答系统能够依据用户提问生成准确答案；文本生成可创作符合人类语言习惯的内容，如新闻报道、故事；信息抽取则负责从文本中提取关键信息，如实体、关系等。这些任务的实现，推动着人机交互向更自然、高效的方向发展。

1.2 自然语言处理的发展历程

自然语言处理的发展经历了多个重要阶段。早期主要采用基于规则

的方法，依赖手工制定规则和模式匹配来处理语言，这种方式虽然简单直观，但面对语言的复杂性和多样性时，表现出明显的局限性，且规则维护成本高昂。

随着机器学习技术的兴起，统计学习方法逐渐成为自然语言处理的主流。该方法通过对大量标注数据的学习，获取语言模式，能够更好地应对语言的不确定性，但对数据质量和数量有着较高要求。

近年来，深度学习技术为自然语言处理带来了重大突破。基于神经网络的模型，如循环神经网络（RNN）、长短期记忆网络（LSTM）、Transformer 架构等，能

【作者简介】蔡丽芸（1986—），女，广东汕头人，本科，高级工程师，研究方向：计算机应用、人工智能应用、信息化项目建设等。

够自动学习语言的特征表示，在各类自然语言处理任务中显著提升了性能，推动该领域进入全新的发展阶段^[3]。

1.3 自然语言处理在计算机应用中的现状

目前，自然语言处理技术在计算机应用中已取得显著成果，并广泛应用于多个领域。在智能客服领域，借助自然语言理解技术，智能客服能够自动响应用户问题，有效提升客户服务的效率与质量；机器翻译技术实现了多种语言之间的自动翻译，尽管仍有提升空间，但已在跨语言交流中发挥重要作用；在社交媒体监测、舆情分析等场景中，文本分类与情感分析技术可快速准确地处理文本，挖掘其中蕴含的信息；智能问答系统，如苹果的 Siri、微软的小娜等，能够依据用户提问生成自然语言回答，提供便捷的信息服务。然而，自然语言处理技术在实际应用中也面临诸多挑战。高质量标注数据是模型训练的基础，但获取成本高，且数据的多样性和平衡性难以保证；深度学习模型虽在性能上有显著提升，但训练和推理效率较低，难以满足部分实时响应场景的需求；自然语言的模糊性、歧义性和多样性，使得机器准确理解和生成自然语言仍存在较大困难。

2 自然语言处理技术的优化策略

2.1 数据预处理的优化

数据是自然语言处理的基石，其质量和数量直接影响模型性能，因此数据预处理阶段的优化至关重要。首先，数据清洗通过去除重复、无关等噪声数据，提升数据的纯净度。其次，数据增强利用同义词替换、句子重组、

回译等手段，扩充数据集规模，增加数据多样性，从而提高模型的泛化能力。再次，采用众包标注、半自动标注等方式进行数据标注，可提高标注效率和准确性，为模型训练提供可靠的数据支撑。最后，针对训练数据中样本数量不均衡的问题，通过数据采样、过采样等方法平衡数据，有助于提升模型的分类性能。

2.2 模型架构的优化

模型作为自然语言处理的核心，优化其结构和参数能够显著提升模型性能^[4]。其一，应用预训练模型，如 BERT、GPT 等，通过在大规模语料上进行预训练，模型可学习到通用的语言特征，再经微调适配具体任务，大幅减少对标注数据的依赖。其二，改进模型架构，将不同的神经网络架构，如 Transformer 与 CNN、RNN 相结合，发挥各自优势，提升模型对自然语言的理解和生成能力。例如，针对 Transformer 架构计算复杂度高的问题，可引入稀疏注意力机制或分块注意力机制进行优化。其三，在资源受限的场景，如移动设备或边缘计算中，设计轻量化模型，运用剪枝、量化等模型压缩技术，在保证性能的前提下减小模型规模。其四，采用多任务学习方式，联合训练多个相关任务，共享模型参数，增强模型对不同任务的适应能力，同时降低训练数据需求。

2.3 算法优化

算法是自然语言处理技术的关键，优化算法能够有效提高模型的训练效率和性能。一方面，采用自适应学习率的梯度优化算法，如 Adam、RMSprop 等，可加快模型收敛速度，提升训练效率。另一方面，通过添加 L1、L2 等正则化项，防止模型过拟合，增强模型的泛化能力。此外，多模态融合将文本与图像、语音等其他模态信息相结合，使模型能够更全面地理解和生成自然语言。例如，在图像描述生成任务中，融合图像和文本特征可生成更准确的描述；在对话系统等需要与环境交互的任务中，运用强化学习算法优化模型的行为策略，能够提升模型的交互性能。

3 自然语言处理技术的实践案例

3.1 智能客服系统

智能客服是自然语言处理技术的重要应用领域，通过优化技术可实现更高效的客户服务。在问题理解与意图识别环节，运用 BERT、Transformer 等深度学习模型对用户问题进行语义理解，提取语义特征后经分类器识别用户意图，如查询订单、投诉建议等。在答案生成与知识检索阶段，依据问题理解结果，借助序列到序列（Seq2Seq）模型或检索知识库的方式，生成准确自然的答案。通过对话管理模块实现多轮对话，根据用户反馈调整回答策略。例如，用户对回答不满意时，主动询问具体需求，优化回答内容，提升用户满意度。

3.2 文本分类

在文本分类任务中，优化策略可从多方面提升模型性能。特征提取与表示方面，利用 Word2Vec、GloVe 等词嵌入技术，将文本转化为低维稠密向量，更好地捕捉语义信息。模型选择与优化时，依据具体任务需求，选用支持向量机（SVM）、深度学习模型等，并对模型进行调整，如使用 BERT 模型编码文本后通过全连接层分类，同时调整 BERT 的层数、

隐藏层大小等超参数。针对训练数据样本不平衡问题，采用 SMOTE 算法对少数类样本过采样，结合同义词替换、句子重组等数据增强技术，平衡数据分布，扩充数据集，提高模型分类性能和泛化能力。

3.3 机器翻译

对于机器翻译，优化措施围绕模型、数据和翻译过程展开。模型架构改进上，以 Transformer 架构为基础，结合多头注意力机制增强对长距离依赖关系的建模能力，通过调整层数、隐藏层大小提升表达能力，并引入稀疏或分块注意力机制优化计算效率。数据增强与多语言学习方面，运用回译、数据拼接等技术扩充训练数据，采用多语言联合训练方式，共享参数，提高模型对不同语言对的适应能力。翻译优化与后处理阶段，利用束搜索等解码算法优化翻译结果生成过程，通过调整束搜索参数平衡准确性和流畅性，对翻译结果进行拼写检查、语法修正等后处理，进一步提升翻译质量^[5]。

4 自然语言处理技术的未来发展趋势

4.1 多模态融合

随着计算机技术的不断发展，多模态数据（如文本、图像、语音等）的融合将成为自然语言处理技术的重要发展方向。通过整合多种模态信息，计算机能够更全面、深入地理解和生成自然语言，在图像描述生成、语音翻译等任务中发挥重要作用，为用户带来更丰富、自然的交互体验。

4.2 预训练模型的进一步发展

预训练模型在自然语言处理领域已展现出强大的性能，未来将朝着更大规模、更高效的方向发展。同时，研究人员会更加关注模型的可解释性和安全性。例如，通过引入知识图谱等结构化知识，增强预训练模型的理解能力，使其在应用中更加可靠、透明。

4.3 强化学习与交互式学习

在对话系统、智能写作等需要与环境交互的任务中，强化学习和交互式学习将发挥关键作用。通过与用户的实时交互，模型能够不断优化自身行为策略，根据用户反馈动态调整输出，提供更符合用户需求的服务，显著提升用户体验。

4.4 轻量化与边缘计算

随着移动设备和边缘计算的普及，自然语言处理技术将更加注重模型的轻量化和效率优化。通过模型压缩、硬件加速等技术，将自然语言处理模型部署到移动设备和边缘设备上，实现低延迟、高效率的自然语言处理应用，满足用户在不同场景下的实时需求。

结语

自然语言处理技术在计算机应用中具有不可或缺的地位，通过对数据预处理、模型架构和算法的优化，能够显著提升该技术的性能和应用效果。本文通过智能客服、文本分类、机器翻译等实际案例，详细展示

了优化后的自然语言处理技术在计算机应用中的实践成效。未来，随着多模态融合、预训练模型发展、强化学习应用以及轻量化技术推进，自然语言处理技术将在更多领域发挥更大作用，为人类社会的发展提供强有力的技术支撑，推动人机交互迈向新的高度。^[8]

引用

[1] 邹洪.基于人工智能的自然语言处理技术在软件测试中的应用研究[J].网络安全和信息化,2024(4):59-61.

[2] 孙羽竹.基于自然语言处理NLP初中人工智能项目式教学实践——以“向网络欺凌Say No”项目为例[J].中国教育技术装备,2024(23):117-120+135.

[3] 李显勇,李齐治,胡容维.“自然语言处理”发展历程的教学探讨[J].电脑知识与技术:学术版,2022,18(26):130-132.

[4] 王睿,张留洋,高志涌,等.面向边缘智能的大模型研究进展[J].计算机研究与发展,2025(62):1-18.

[5] 孙磊.基于人工智能的自然语言处理技术分析[J].集成电路应用,2025,42(2):146-147.

