



浙江工业大学

《文献检索与论文写作》 课程报告

题目： 基于高频关键词的大语言模型研究热点
及发展趋势探析

组 长： 202103150503 陈王子
组 员： 202105080116 王之楚
： 202103151422 温家伟
专 业： 大数据科学与技术专业
班 级： 大数据分析方向
提交日期： 2024 年 6 月

摘要

近年来，大语言模型（LLMs）的快速发展已成为自然语言处理（NLP）和人工智能（AI）领域内引人瞩目的焦点。这一趋势的兴起，始于深度学习和神经网络语言模型的成熟与广泛应用，标志着自然语言理解和生成技术的重大飞跃。本文综合分析了 2018 年至 2024 年间，以“大语言模型”为主题，源自学术期刊的 72 篇精选文献。通过 Python 对数据文献进行高频关键词词频统计、高频关键词相似度矩阵构建以及关键词层次聚类分析，探究大语言模型领域的前沿研究热点，并对未来发展方向进行展望。研究发现，提升模型的通用性与智能水平，实现通用人工智能（AGI），优化神经网络框架，提升模型理解与生成能力，跨领域知识迁移以及多语言信息处理等成为了当今大语言模型技术的研究热点。根据数据分析结果，优化模型的规模与效能、多模态信息处理、社会伦理与隐私保护、应用领域的深度融合与定制化开发、规模效应和工作机理的深入探索成为未来大语言模型的几个关键发展方向和趋势。

关键词：大语言模型，LLM，研究热点，趋势分析，数据可视化

Research hotspots and development trends of large language models based on high-frequency keywords

ABSTRACT

In recent years, the rapid development of Large Language Models (LLMs) has become a remarkable focal point in the fields of Natural Language Processing (NLP) and Artificial Intelligence (AI). This surge, rooted in the maturation and widespread application of deep learning and neural network language models, signifies a significant leap forward in the technology of natural language understanding and generation. This paper comprehensively analyzes 37 selected academic journal articles from 2018 to 2024, all centered on the theme of 'Large Language Models'. Utilizing Python for high-frequency keyword frequency statistics, construction of a high-frequency keyword similarity matrix, and hierarchical clustering analysis of keywords, this study delves into the cutting-edge research hotspots within the domain of large language models and projects future development directions. The findings reveal that enhancing model generality and intelligence levels towards achieving Artificial General Intelligence (AGI), optimizing neural network frameworks, improving model understanding and generation capabilities, cross-domain knowledge transfer, and multilingual information processing have emerged as current research hotspots in the field of large language model technology. Based on the data analysis results, the optimization of model scale and efficiency, multimodal information processing, social ethics and privacy protection, in-depth integration and customized development in application domains, and exploration of the scale effects and inner workings of models are identified as key future development trends for large language models.

KEY WORDS: Large Language Models, LLM, Research Hotspots, Trend Analysis, Data Visualization

目 录

摘要.....	1
ABSTRACT.....	2
引言.....	4
第一章 综述.....	5
1.1 数据来源.....	5
1.2 论文分类.....	5
1.3 国内外来源.....	6
1.4 时间趋势.....	6
1.5 研究方法和数据处理工具.....	7
第二章 关键词分析.....	8
2.1 高频关键词词频统计.....	8
2.2 高频关键词相似度矩阵分析.....	9
2.3 关键词聚类分析.....	10
第三章 当前研究热点.....	12
3.1 提升模型的通用性与智能水平.....	12
3.2 实现通用人工智能（AGI）.....	12
3.3 优化神经网络框架.....	12
3.4 提升模型理解与生成能力.....	12
3.5 跨领域知识迁移以及多语言信息处理.....	13
第四章 未来发展方向.....	14
1. 优化模型的规模与效能.....	14
参考文献.....	16

引言

在过去的数十年间，信息技术的迅猛发展催生了海量文本数据，覆盖了从社交媒体交流、新闻报道到学术文献和网络文本的广泛领域。而大语言模型作为这一时代背景下的新兴力量，正以前所未有的速度吸收并学习着这些多元化的信息资源。随着人工智能的不断演进，大语言模型领域正经历着深刻的变革，其核心在于将海量数据、海量算力与自然语言处理的深度融合，以此推动语言理解和生成能力的边界拓展。

大语言模型的应用范围横跨数据的采集、整合、分析至解释和应用的全链条。初始阶段，通过互联网爬虫、文本挖掘工具等手段，各类文本数据被高效收集；随后，借助于云计算平台和先进的数据管理技术，这些数据得以有序存储和维护。在此基础上，运用深度学习框架和自然语言处理算法，对海量文本进行深度剖析，挖掘其中隐含的知识结构和语义关联，从而为语义理解、智能对话和内容生成等应用场景提供坚实的基础。最后，通过自然语言生成和可视化技术，将复杂的分析结果转化为易于理解的形式，促进人机交互和决策制定的优化。

本研究致力于通过全面回顾与剖析，洞察大语言模型领域的前沿进展与实践现状。面对数据量的激增，如何有效利用这些资源以提升语言模型的性能，成为理论研究与实际应用共同追求的目标。通过对过去五年来围绕“大语言模型”主题的文献进行系统梳理，本研究不仅揭示了范式革新在语言理解、文本生成等方面的作用，同时也关注了全球范围内该领域的研究动态，展示了技术迭代的国际视野。

通过 Python 实施的关键词分析，为识别研究趋势与未来导向提供了实证支持。例如，分析表明，“大语言模型”、“ChatGPT”、“人工智能”、“AI 大模型”、“通用人工智能”等概念近年来频繁出现，彰显了它们在推动大语言模型进步中的核心角色。与此同时，数据隐私、版权归属和算法偏见等议题也日益受到重视，凸显了在技术创新过程中平衡伦理考量的重要性。

综上所述，本文通过跨学科、跨地域的文献综述，剖析了大语言模型技术领域的关键议题，为理解其演变历程和未来前景提供了全景式的视角。展望未来，伴随技术成熟和规范健全，大语言模型有望在更广阔的场景下展现其价值，引领自然语言处理领域的智能化、个性化和多元化发展。

第一章 综述

1.1 数据来源

研究基于中国知网期刊全文数据库为文献数据来源，设置检索主题为“大数据医疗”，文献来源设置为学术期刊、学术论文及会议，检索时间为 2018 年至 2024 年。对检索到的文献进行甄别筛选，最后选取了 72 篇文献作为本次研究的分析数据。

1.2 论文分类

我们可以将这些文献按照不同的侧重和研究焦点进行分类。《国内外大语言模型的图书情报应用探讨》^[3]和《SikuBERT 与 SikuRoBERTa: 面向数字人文的《四库全书》预训练模型构建及应用研究》^[8]分别从技术进展、行业应用和古文智能处理的角度探讨了预训练模型的影响。《GPT 系列大语言模型在自然语言处理任务中的鲁棒性》^[2]研究了 GPT 系列模型在自然语言处理任务上的鲁棒性，揭示了模型在特定任务中的优势和局限。《面向文本推理的知识增强预训练语言模型》^[14]介绍了一种利用知识图谱增强预训练模型的方法，以提升文本推理能力。《神经机器翻译综述》^[4]全面概述了 NMT 的发展，从模型架构到神经网络应用，以及面临的挑战。《大模型时代的自然语言处理:挑战、机遇与发展》^[1]和《ChatGPT 大模型技术发展与应用》^[13]分别探讨了 NLP 在更广阔社会经济环境中的作用和 ChatGPT 的里程碑式影响。《AI 大模型发展综述》^[17]概述了 AI 大模型的发展历程，特别是以 BERT 和 GPT 为代表的模型，探讨了它们的技术核心和原理，以及对 AI 领域的深远影响。

《Parameter-efficient fine-tuning of large-scale pre-trained language models》^[18]探讨了大规模预训练语言模型参数高效微调的方法，针对模型规模化后如何有效地适应特定下游任务这一研究议题，尤其关注在任务特定设置下微调超越上下文学习的性能。《Large language models for reducing clinicians' documentation burden》^[21]介绍了一种基于 GPT-4 的临床总结方法，该方法可以通过大型语言模型减轻临床医生的记录负担。然而，在高优先级任务上的前瞻性评估将是测试其潜力的真正考验。

1.3 国内外来源

在大语言模型的研究领域，我们观察到了显著的全球协作和跨学科研究趋势，这从所选文献的来源可见一斑，它们既包括了国内顶尖学术机构和知名期刊，也覆盖了国际权威学术出版平台。

在国内，参与研究的机构有清华大学、中国人民大学、复旦大学、华东师范大学、哈尔滨工业大学、苏州大学、南京邮电大学、中国科学院软件研究所、北京市信息服务工程重点实验室等。和发布成果的期刊有《中文信息学报》、《软件学报》、《计算机学报》、《计算机科学》、《中国科学》、《自动化学报》、《图书馆理论与实践》、《计算机研究与发展》、《小型微型计算机系统》、《工程科学学报》等。这些研究深入探讨了模型训练数据的集成、模型评估标准、自然语言处理系统设计、大规模语言模型平台建设的挑战等方面，彰显了我国学者在推动大语言模型技术进步和应用创新上的广度和深度。

而在国际舞台上，相关研究被发表在诸如《ACM Computing Surveys》、《Nature Machine Intelligence》、《ACL (Association for Computational Linguistics)》、《IEEE Access》、《EMNLP (Conference on Empirical Methods in Natural Language Processing)》、《NAACL (North American Chapter of the Association for Computational Linguistics)》等具有高影响力的国际期刊和会议论文集上。这些研究聚焦于大语言模型在不同应用场景下的表现优化、基于大数据的文本分类和情感分析、模糊逻辑在文本理解中的应用、深度学习架构在自然语言生成中的创新、分布式计算框架下的模型训练策略、人机协同在高级文本分析中的融合，以及利用大模型识别复杂语言模式的能力等前沿议题。这反映出全球研究者对大语言模型技术在提升自然语言理解和生成精度、促进个性化交流及自动化信息处理等方面的浓厚兴趣和积极贡献。

1.4 时间趋势

通过对所选文献的时间序列分析，我们发现 2022 年至 2024 年间，大语言模型领域的文献产出最为丰富，伴随着年初 ChatGPT 的出现迎来了井喷，研究主题广泛多样。相比之下，2021 年前的相关研究则相对较少，当时自然语言处理领域并不把预训练的大语言模型作为主流范式，而是主要聚焦于理论构想与初步验证阶段。这一早期阶段的研究着重于探索大语言模型的潜在应用场景，以及它们可能引发的技术革新与挑战，为后续的实践探索和技术迭代打下了坚实的理论

基石。

2022 年后，伴随着一些早期大模型的成功推出，期刊的研究深度和针对性显著增强，更多地关注于提升技术性能、解决实际问题，尤其是在优化计算效率、保障数据隐私、促进多模态数据融合等方面。同时，针对特定任务和场景的大语言模型应用研究日益深化，推动了大语言模型在实际场景中的落地实施。而 2021 年前，“力大砖飞”式的堆算力、堆数据并不是业内的 SOTA 主流，这表明，随着技术的逐步成熟，研究重心正向解决现实世界中的技术难题和改进现有解决方案转移。

根据文献年代分布的结果，大语言模型技术的发展前景极其广阔。其快速演进更是彰显了在自然语言处理领域的实用价值和战略重要性，为智能文本分析提供了更为丰富的数据视角，摘下人工智能王冠上的明珠。

1.5 研究方法和数据处理工具

本次研究运用 Python 及其相应的软件包对大语言模型领域的文献资料进行深度分析，主要通过关键词频率统计、关键词相似性评估及层级聚类分析，以期提炼出该领域的核心研究趋势与未来探索方向。

我们首先将几十篇论文的文本数据导入 Python，随后进行预处理，包括去除噪声和无关信息，确保分析的有效性。然后，我们采用 TF-IDF 算法从每篇文献中抽取关键术语，统计这些术语在整个文献集中的出现频率，进而形成高频关键词清单。借助 Matplotlib 可视化工具，我们将绘制关键词频率曲线图，直观展现关键词的分布情况。之后，我们利用 Jaccard 相似系数量化高频关键词之间的关联程度，构建关键词相似度矩阵，通过图形化方式呈现关键词间的亲疏关系。最后，我们将关键词相似度矩阵作为输入，执行层级聚类算法，生成聚类矩阵图，旨在揭示关键词间深层次的结构性联系。

第二章 关键词分析

关键词分析是对文献中的关键词进行提取、聚类 and 可视化分析等。通过对文献关键词的提取和分析，可以快速概括文献的主题和内容，为研究者提供对文献的快速理解，并且关键词分析能够帮助研究者发现文献中涉及的研究热点和前沿领域，有助于把握学术动态和趋势。

下面是对大语言模型文献关键词分析的详细过程。

2.1 高频关键词词频统计

在进行关键词分析之前，首先找出大语言模型论文的摘要部分，并提取其中的关键词。这些关键词通常能够较好地概括文献的主题和内容，用于后续分析。

Echarts 是一个强大的数据可视化库，可以用来制作交互式的图表，便于用户进行数据探索和分析。运用 echarts 将关键词词频统计的结果可视化为饼状图，以直观地展示关键词的分布情况。

大语言模型文献的关键词词频统计的可视化饼状图如下：

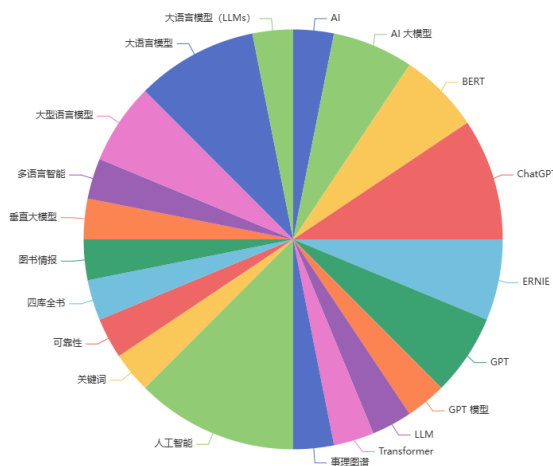


图 2-1 关键词词频统计图 Keyword word frequency statistics chart

可以看出，“大语言模型”、“ChatGpt”、“人工智能”、“AI 大模型”、“BERT”、“ERNIE”等关键词出现的频率较高。

“大语言模型”和“ChatGPT”在文献中频繁出现，反映了对于大型语言模

型研究和探索的焦点。同时，“ChatGPT”在自然语言处理领域的对话生成和对话系统方面有着重要的应用和研究，反映出大语言模型在不同场景中的应用和影响。“BERT”是一种广泛应用的预训练语言模型，该词的出现频率较高也反映了该领域的技术趋势和发展方向。“ERNIE”是百度提出的一种基于持续学习的、通用的语义理解预训练框架，融合多种知识和语义表示，相比于传统的预训练模型（如 BERT），提升了模型对于自然语言理解的能力，该高频词反映了大模型领域新的方法和技术创新。“人工智能”和“AI 大模型”是探讨大型语言模型在人工智能领域中应用和发展的相关理论和方法的文献中常见的关键词，反映出该领域中的理论探索。

综合来看，高频关键词在一定程度上反映了大语言模型领域的研究热点、技术趋势、理论探索以及应用场景等，对于了解该领域的发展动态和关注重点具有一定的指导意义。

2.2 高频关键词相似度矩阵分析

使用 Word2Vec 词嵌入模型对文献中提取的关键词进行向量化处理，计算它们之间的相似度。词嵌入模型能够将关键词映射到高维空间中的向量，相似的词在向量空间中距离较近，不相似的词则距离较远。本文运用 python 编程实现 Word2Vec 模型，得到的关键词相似度矩阵如下图所示：

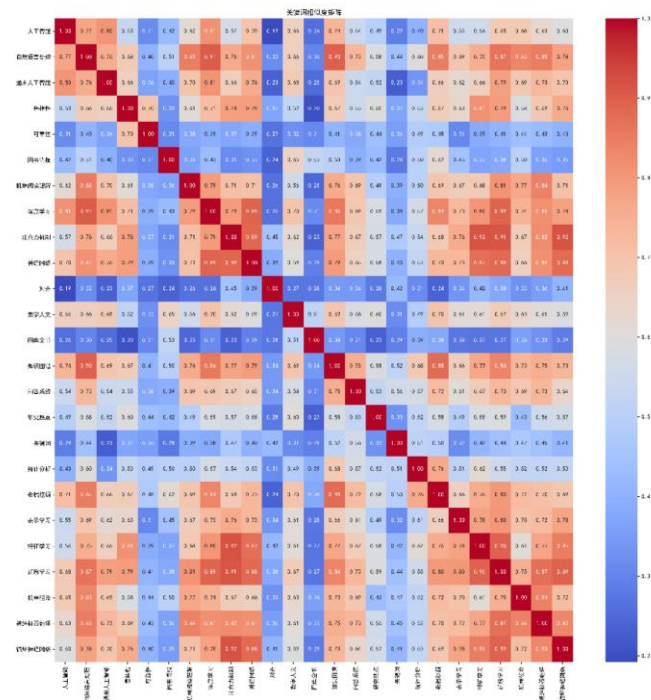


图 2-2 关键词相似度矩阵 Keyword similarity matrix

根据大语言模型关键词相似度矩阵可以看出,“人工智能”、“自然语言处理”、“通用人工智能”相关性较高。自然语言处理是人工智能的一个子领域,旨在让计算机能够理解、解释、操作人类语言。通用人工智能则是一种对于普遍任务具有广泛适应性和类人水平智能的概念,是人工智能领域的一个追求目标。因此,这三个关键词之间具有较高的相关性。

“机器阅读理解”、“深度学习”和“神经网络”相关性也较高。机器阅读理解是利用计算机技术让机器能够理解和回答文本信息的过程,深度学习则是一种机器学习方法,通过构建深层神经网络来实现对数据的学习和表示,而神经网络是深度学习的核心。这三个关键词的相关性表明在处理自然语言理解任务时,深度学习和神经网络是常用的技术手段,在机器阅读理解领域中发挥着重要作用。

“数据挖掘”、“表示学习”、“迁移学习”、“机器翻译”和“循环神经网络”的相关性较高。数据挖掘是从大规模数据中提取有价值信息的过程,表示学习、迁移学习和循环神经网络都是机器学习方法,机器翻译是实现不同语言间自动翻译的过程。说明它们在处理大规模数据和自然语言任务中有很高的相关性,特别是在涉及到序列数据处理应用和知识迁移方面。

通过对这些关键词相关性的深入分析可以帮助我们更好地理解人工智能领域的发展趋势、技术交叉点和研究热点,有利于对该类文献的理解和应用。

2.3 关键词聚类分析

利用计算得到的词向量相似度,可以进行层次聚类。层次聚类是一种将数据点逐渐合并成越来越大的簇的方法,最终形成一个层级结构的聚类结果。这可以帮助我们发现关键词之间的内在结构和关联性。

本文使用 Python 中的可视化库来绘制层次聚类的树状图,展示关键词之间的层级关系和聚类结果。这样的可视化能够帮助我们更直观地理解关键词之间的关联性和组织结构。

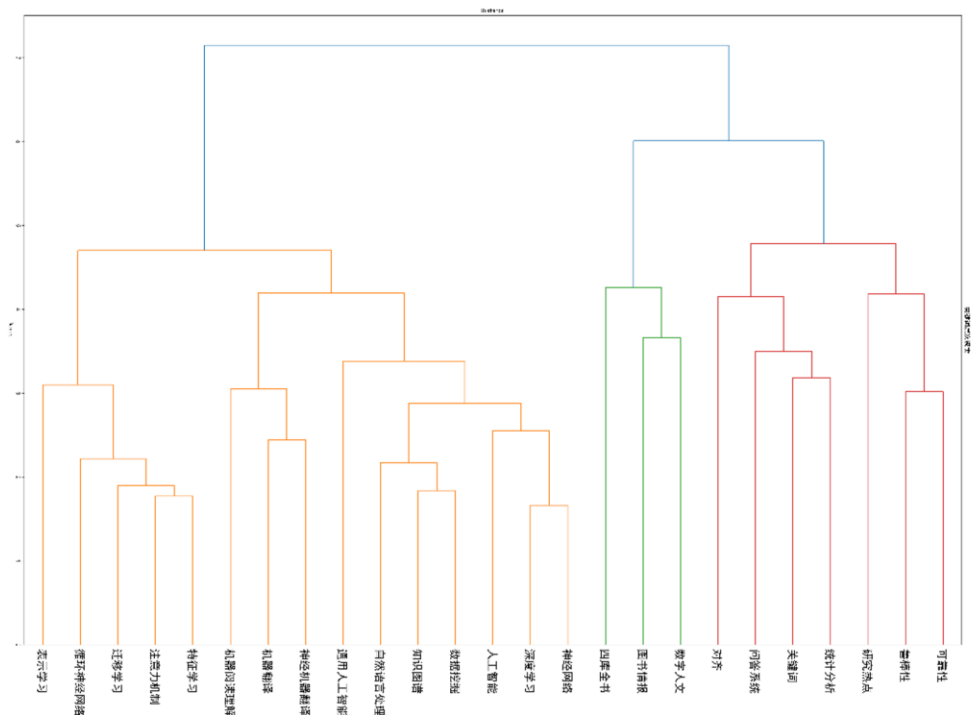


图 2-3 关键词聚类图 Key words Cluster graph

这些关键词在聚类中呈现出不同的特征和关联，通过上图，我们可以看到这些关键词大致聚为三类：

第一类是“可靠性”、“鲁棒性”、“研究热点”、“统计分析”、“问答系统”等关键词。这些关键词代表了对系统性能和数据质量的关注。可靠性和鲁棒性是指系统在面对异常情况或噪声时的稳定性和表现能力。统计分析和问答系统表示了语言模型的应用领域和研究热点。

第二类是“数字人文”、“图书情报”、“四库全书”等关键词。这些关键词表明可以利用数字技术进行人文学科研究和实践，对图书馆、档案馆等信息机构的管理和服务。这些关键词反映了大语言模型在人文社科领域中数字化、信息化和文献资源管理的作用。

第三类是“神经网络”、“深度学习”、“人工智能”、“知识图谱”、“自然语言处理”、“机器翻译”、“特征学习”等关键词。这些关键词表明了人工智能和机器学习领域的核心技术和应用，反映了人工智能和机器学习领域的核心技术和研究方向，它们之间存在着密切的关联和相互作用。

通过对这些聚类的分析，我们看到了不同领域和主题之间的关联和差异，这有助于我们更好地理解相关研究领域的发展动态和关键问题，为进一步研究和应用提供有益的参考。

第三章 当前研究热点

3.1 提升模型的通用性与智能水平

在大语言模型的前沿探索中，核心议题之一便是增强其通用性和智能层次。早期模型往往参数较少，表现不尽如人意。^[12]尽管近年来模型容量的膨胀显著提升了其在自然语言处理任务中的效能，模型的鲁棒性、泛化能力以及对复杂上下文的理解仍存局限^[2]。为解决这些问题，为此，研究者正致力于解码语义关联与逻辑构造^[7]，让大语言模型发展出自提示的逻辑链（CoT），从而提升其在复杂场景下的表现^[24]。

3.2 实现通用人工智能（AGI）

追求通用人工智能（AGI）是大语言模型研究的长远目标和终极愿景。尽管目前的模型在特定任务上表现出色、在限定领域展现出了远超人类的能力，但要达到人类水平的智能，还需跨越理论和技术上的重重障碍。研究者正从不同角度探索 AGI 的路径，包括强化预训练策略^[10]、改进模型架构和设计更高效的算法^{[6][17]}。这些努力均为了构建兼具深度理解和广泛适应性的智能体（Agent），使其在纷繁复杂的现实世界中仍能游刃有余地处理人类的命令。

3.3 优化神经网络框架

神经网络框架的优化，是大语言模型性能飞跃的关键推手。自词向量时代^[16]、到神经机器翻译兴起至深度学习的蓬勃演进^[9]，再到 Transformer 架构的横空出世^[25]，该领域持续涌现的创新成果彰显了技术迭代的活力^[4]。近期的研究重点放在了如何减少模型参数数量的同时保持或提升性能^[23]，以及如何利用知识增强来改善文本推理能力^[18]。

3.4 提升模型理解与生成能力

大语言模型的理解与生成能力的强化，构成了研究领域的又一重心^[5]。这不

仅涉及到改进模型的自然语言处理能力，还包括增强其在代码生成^[20]、财会和法律文书^[24]、医疗文档处理和精准肿瘤学^{[21] [22]}等领域的应用。通过持续优化，研究人员致力于让模型能够更加准确地理解和生成高质量的文本，为人类社会的多元化需求提供有力支撑。

3.5 跨领域知识迁移以及多语言信息处理

跨领域知识迁移与多语言信息处理，是大语言模型研究不可或缺的一环^[11]。随着全球化的加深，多语种兼容性已成为评判模型优劣的重要标准^[15]。同时，跨领域知识的融合有助于模型在不同行业和应用场景中发挥更大的作用，如图书情报学^[3]和数字人文^[8]等领域。跨领域的大模型似乎可以构建一个无界的知识生态，使模型成为连接人类文明各分支的智慧桥梁^[19]。

综上所述，大语言模型的研究热点覆盖了从基础理论到应用实践的全方位探索，旨在提升模型的综合性能，为人类社会的数字化转型注入不竭动力，最终朝着实现真正的通用人工智能迈进。

第四章 未来发展方向

预训练大语言模型的迅猛演进，已然显著重塑了自然语言处理（NLP）领域的格局，其影响深远且前景无限。从语义理解的精进、个性化交互体验的提升，到信息检索与内容生成的效率优化，NLP 技术与各行各业的融合正开辟出崭新的应用疆域。

1. 优化模型的规模与效能

优化大语言模型的规模与效能是未来发展的基石。面对日益增长的数据量和计算需求，研究者致力于探索更高效、更经济的模型训练方法。这包括但不限于精简模型架构，采用参数效率更高的微调策略[18]，以及优化模型压缩技术，确保在不牺牲性能的前提下减小模型体积，甚至集成进手机[14]。此外，随着硬件技术的进步，如何充分利用 GPU、TPU 等加速器资源，以更低的成本实现模型的快速迭代与部署，降低模型的训练门槛，进行分布式或者敏捷训练，亦是研究的热点。

2. 多模态信息处理

多模态信息处理是大语言模型拓展感知世界能力的关键。传统的语言模型主要处理文本信息，而未来的发展趋势将融合视觉、听觉等多种感官输入，实现对复杂现实环境的全面理解[13]。例如，结合图像和视频数据，使模型能够理解场景语境，增强其情境感知能力；整合语音信号，提升模型的口语理解和生成能力。多模态融合不仅丰富了模型的信息输入维度，更是横跨计算机视觉和计算机图形学，为其在虚拟助手、AI 绘画、自动驾驶、智能家居等应用场景中提供了更广阔的发展空间。

3. 社会伦理与隐私保护

随着大语言模型在社会各领域的广泛应用，社会伦理与隐私保护问题日益凸显。如何在保障用户隐私安全的同时，确保模型输出的公正性、无偏见，避免传播错误信息或歧视性言论，成为亟待解决的问题[11]。研究者正积极探索建立模型伦理准则，引入可解释性机制，使模型决策过程透明化，同时加强数据脱敏处理，保护个人隐私，构建更加负责任的人工智能生态。

4. 应用领域的深度融合与定制化开发

大语言模型的应用领域深度融合与定制化开发是其商业价值和社会影响力的体现。针对特定行业的需求，如医疗健康、法律咨询、教育科技等，定制化的大语言模型能够提供更加精准的服务[21] [22]。这不仅需要模型具备强大的领域知识，还需要在数据集构建、模型训练和评估指标等方面进行专门设计，以适应特定领域的语言风格和专业术语，建立通用语言模型的特化知识库，实现从通用

智能向专业智能的转变。

5. 规模效应和工作机理的深入探索

大语言模型的规模效应和工作机理是当前研究的核心议题之一。虽然大规模模型展现出卓越的性能，但其内在的工作机制仍存在诸多未解之谜。研究者正致力于揭示模型内部的运作规律，探索其如何在高维空间中捕捉语言的复杂结构，以及如何通过自监督学习策略从海量数据中自动提取知识[7]。对模型工作机理的深入了解，将协同计算机科学、数学、神经科学、哲学等基础学科，将为后续模型设计提供理论指导，让人类对智能的涌现产生全新的见解。

综上所述，未来大语言模型的发展将围绕模型效能提升、多模态信息处理、社会伦理考量、应用领域拓展以及理论机制探索等多个维度展开，构建更加智能、可靠、多元和负责任的人工智能体系，于今日的地平线上一窥通用人工智能的曙光。

参考文献

- [1] 车万翔, 窦志成, 冯岩松. 大模型时代的自然语言处理挑战、机遇与发展[J]. 中国科学信息科学, 2023, 53(09): 1645-1687.
- [2] 陈炫婷, 叶俊杰, 祖璨. GPT 系列大语言模型在自然语言处理任务中的鲁棒性[J]. 计算机研究与发展, 2024, 61(05): 1128-1142.
- [3] 洪贇, 叶鹰, 佟彤. 国内外大语言模型的图书情报应用探讨[J]. 图书馆理论与实践, 2024(02): 72-80. DOI: 10.14064/j.cnki.issn1005-8214.2024.02.015.
- [4] 李亚超, 熊德意, 张民. 神经机器翻译综述[J]. 计算机学报, 2018, 41(12): 2734-2755.
- [5] 李舟军, 范宇, 吴贤杰. 面向自然语言处理的预训练技术研究综述[J]. 计算机科学, 2020, 47(03): 162-173.
- [6] 李舟军, 王昌宝. 基于深度学习的机器阅读理解综述[J]. 计算机科学, 2019, 46(07): 7-12.
- [7] 舒文韬, 李睿潇, 孙天祥. 大型语言模型: 原理、实现与发展[J]. 计算机研究与发展, 2024, 61(02): 351-361.
- [8] 王东波, 刘畅, 朱子赫. SikuBERT 与 SikuRoBERTa: 面向数字人文的《四库全书》预训练模型构建及应用研究[J]. 图书馆论坛, 2022, 42(06): 31-43.
- [9] 王乃钰, 叶育鑫, 刘露. 基于深度学习的语言模型研究进展[J]. 软件学报, 2021, 32(04): 1082-1115.
- [10] 王鑫雷, 李帅驰, 杨志豪. 基于预训练语言模型的中文知识图谱问答系统[J]. 山西大学学报(自然科学版), 2020, 43(04): 955-962.
- [11] 魏子舒, 韩越, 刘思浩. 2021 至 2023 年人工智能领域研究热点分析述评与展望[J]. 计算机研究与发展, 2024, 61(05): 1261-1275.
- [12] 奚雪峰, 周国栋. 面向自然语言处理的深度学习研究[J]. 自动化学报, 2016, 42(10): 1445-1465.
- [13] 夏润泽, 李丕绩. ChatGPT 大模型技术发展与应用[J]. 数据采集与处理, 2023, 38(05): 1017-1034.
- [14] 熊凯, 杜理, 丁效. 面向文本推理的知识增强预训练语言模型[J]. 中文信息学报, 2022, 36(12): 27-35.
- [15] 徐月梅, 胡玲, 赵佳艺. 大语言模型与多语言智能的研究进展与启示[J]. 计算机应用, 2023, 43(S2): 1-8.
- [16] 郁可人, 傅云斌, 董启文. 基于神经网络语言模型的分布式词向量研究进展[J]. 华东师范大学学报(自然科学版), 2017(05): 52-65+79.
- [17] 张乾君. AI 大模型发展综述[J]. 通信技术, 2023, 56(03): 255-262.

- [18] Ding, N., Qin, Y., Yang, G . Parameter-efficient fine-tuning of large-scale pre-trained language models[J]. Nature Machine Intelligence, 2023, 5(March): 220-235.
- [19] Bonan Min, Hayley Ross, Elior Sulem. Recent Advances in Natural Language Processing via Large Pre-Trained Language Models: A Survey[J]. ACM Computing Surveys, 2024, 57(2): 1-38.
- [20] Yeo, Sangyeop, Yu-Seung Ma, Sang Cheol Kim. “Framework for evaluating code generation ability of large language models[J]. ETRI Journal, 2024, 46(1): 106-117.
- [21] Roberts, K. Large language models for reducing clinicians’ documentation burden[J]. Nature Medicine, 2024, 30(4): 942-943.
- [22] Truhn, D., Eckardt, JN., Ferber, D. Large language models and multimodal foundation models for precision oncology[J]. npj Precision Oncology, 2024, 8(1): 72.
- [23] Hoffmann, J., Borgeaud. (2024). Training compute-optimal large language models[J]. Proceedings of the 36th International Conference on Neural Information Processing Systems, 30016–30030.
- [24] Yang, X., Wang, Z., Wang, Q. (2024). Large language models for automated Q&A involving legal documents: a survey on algorithms, frameworks and applications[J]. International Journal of Web Information Systems,13-14
- [25] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J. Attention Is All You Need[J]. Advances in Neural Information Processing Systems 30 (NIPS 2017), 2017.