

人工智能生成内容中的话语认知分析

沈 越, 赵振强

浙江越秀外国语学院应用外语学院, 浙江 绍兴

收稿日期: 2024年8月12日; 录用日期: 2024年9月4日; 发布日期: 2024年9月12日

摘要

本研究探讨了人工智能(AI)生成内容中的话语认知特征, 旨在分析这些内容在语言表达、认知负荷、语义连贯性等方面的特点与差异。通过系统文献回顾, 概述了AI生成模型的发展历程和技术基础, 特别是对GPT-3等先进模型进行了深入探讨。研究采用质性与量化分析相结合的方法, 从语言复杂度、情感色彩和认知负荷等多维度, 对比分析了AI生成文本与人类生成文本。实验结果显示, AI生成内容在语言一致性和逻辑连贯性方面表现优秀, 但在语义深度和情感表达上存在一定局限。进一步分析表明, 这些差异可能源于算法设计、训练数据和生成机制的不同。本文提出了提升AI生成内容质量的建议, 并为未来人工智能语言模型的优化和应用提供了科学依据和参考。

关键词

人工智能, 话语认知, 语言表达, 认知负荷, 语义连贯性

Cognitive Analysis of Discourse in AI-Generated Content

Yue Shen, Zhenqiang Zhao

School of Applied Foreign Languages, Zhejiang Yuexiu University, Shaoxing Zhejiang

Received: Aug. 12th, 2024; accepted: Sep. 4th, 2024; published: Sep. 12th, 2024

Abstract

This study explores the cognitive features of discourse in content generated by Artificial Intelligence (AI), aiming to analyze the characteristics and differences in terms of language expression, cognitive load, and semantic coherence. Through a systematic literature review, the development and technical foundations of AI-generated models are outlined, with an in-depth discussion of advanced models such as GPT-3. Employing a combination of qualitative and quantitative analysis methods, the study compares AI-generated texts with human-generated texts across multiple dimensions,

including language complexity, emotional tone, and cognitive load. Experimental results indicate that AI-generated content performs well in linguistic consistency and logical coherence, yet shows certain limitations in semantic depth and emotional expression. Further analysis suggests these differences may stem from variations in algorithm design, training data, and generation mechanisms. The paper proposes recommendations for improving the quality of AI-generated content and provides scientific evidence and references for the optimization and application of future AI language models.

Keywords

Artificial Intelligence, Discourse Cognition, Language Expression, Cognitive Load, Semantic Coherence

Copyright © 2024 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

在当今飞速发展的信息化时代,人工智能(AI)技术在各个领域展现出巨大的潜力,尤其是在自然语言处理和生成方面,以GPT-3为代表的新一代语言模型生成的内容已经具备相当高的质量和复杂性。然而, AI生成内容在话语认知层面的特征和表现仍引发广泛关注与讨论。为了更深入地理解这一现象,本研究旨在分析AI生成内容中的话语认知特征,探讨其在语言表达、语义连贯性和认知负荷等方面的表现差异及其潜在原因。随着AI技术的进步,机器生成的文本在很多应用场景中扮演着越来越重要的角色,但其与人类生成内容在认知和语言特征上的差异尚未得到充分研究。本研究的目的是系统性地分析AI生成内容在这些方面的特性,旨在揭示其在话语一致性、逻辑性和情感表达方面的优劣,从而为未来AI语言模型的优化提供科学依据。

2. 文献综述

2.1. 人工智能生成内容的概述

2.1.1. 人工智能技术的发展

人工智能(AI)技术的发展历史悠久,自20世纪50年代以来经历了多个阶段和转折点。早期的人工智能主要依赖符号主义方法,即通过明确的规则和逻辑推理来模拟人类智能[1]。然而,这种方法由于需要大量手工编写规则,在处理复杂和动态的问题时效率低下,逐渐被更为灵活和自适应的数据驱动方法所取代。

20世纪80年代和90年代,专家系统作为人工智能的重要应用取得了一定的成功。然而,这一时期的人工智能技术在应对实际问题时仍面临许多挑战,主要表现为知识获取瓶颈和系统的可扩展性问题[2]。随着计算机性能的不断提升和大数据的广泛应用,机器学习,特别是深度学习技术在21世纪初得到了迅速发展。深度学习通过构建多层神经网络,能够自动从大量数据中提取复杂特征,从而在语音识别、图像识别和自然语言处理等领域取得了突破性进展[3]。

2.1.2. 内容生成算法(如GPT-3、DALL-E等)

现代人工智能在内容生成方面的成就尤为显著,其中GPT-3和DALL-E是两项代表性的突破性技

术。GPT-3 (Generative Pre-Trained Transformer 3)是由 OpenAI 开发的自然语言生成模型，它基于 Transformer 架构，经过大规模语料库的预训练，能够生成具有高语法和语义一致性的文本[4]。GPT-3 的参数数量达到 1750 亿，使其能够在多种任务中展示出接近人类水平的文本生成能力，包括对话、翻译、写作和代码生成等应用。

同时，DALL-E 作为图像生成领域的代表性模型，也由 OpenAI 开发。DALL-E 结合了生成对抗网络 (GAN) 和 Transformer 架构，能够根据文本描述生成逼真的图像[5]。例如，用户输入“一个戴红色帽子的黄色鸟”，DALL-E 可以生成多种符合描述的图像样本。这样的能力展示了文本与视觉内容生成技术融合的巨大潜力。

2.2. 话语分析理论

2.2.1. 话语分析的基本概念

话语分析(Discourse Analysis)作为研究语言使用和交际的重要理论，关注语言在社会情境中的实际使用。话语分析不仅仅研究个体句子的结构和意义，更关注语言在上下文中的连贯性、功能和社会意义[6]。其核心概念包括连贯性、衔接、语境、话语模式等。

连贯性和衔接是话语分析中的基本概念。连贯性指的是话语在语义上的一致性和逻辑关系，而衔接则涉及用明示的语言手段(如代词、连接词)来体现话语中的连贯性[7]。语境包括话语发生的社会、文化环境以及参与者的背景信息，是理解话语意义的关键。

2.2.2. 话语分析的模型与方法

话语分析的方法和模型众多，其中系统功能语言学(Systemic Functional Linguistics, SFL)和批判话语分析(Critical Discourse Analysis, CDA)是两种主要的分析框架。SFL 由 Halliday 提出，强调语言的功能性，认为语言不仅是传递信息的工具，更是在社会互动中实现特定功能的手段[8]。SFL 通过分析语法、语篇和语境等层次，揭示语言如何在不同情境中实现意义的构建。

批判话语分析(CDA)由 Fairclough 等人发展起来，旨在揭示语言与权力、意识形态之间的关系。CDA 认为语言不仅是中立的交流工具，也是反映和维护社会不平等的手段，因而需要通过批判性视角进行分析[9]。CDA 的方法包括文本分析、语篇实践和社会实践三层面，揭示语言在维护和挑战社会权力关系中的作用。

2.3. 话语认知的相关研究

2.3.1. 认知语言学基础

认知语言学(Cognitive Linguistics)是研究语言与认知过程之间关系的理论，强调语言反映了人类的思维方式和世界观。认知语言学的重要概念包括框架(Frames)、概念隐喻(Conceptual Metaphors)和心理空间(Mental Spaces)等[10]。这些理论工具用于解释语言如何通过认知过程构建和传达复杂意义。

框架理论(Frame Theory)强调人们通过既有的知识结构(即框架)来理解新信息。每个框架包含了一组相关的概念和关系，用于组织和解释现实世界中的现象[11]。例如，“购买”这一框架包含了买家、卖家、商品、交易过程等相关元素。

概念隐喻理论(Conceptual Metaphor Theory)则揭示了抽象概念如何通过具体经验和感知来理解。Lakoff 和 Johnson 在其经典著作《我们赖以生存的隐喻》中，提出了诸如“时间是金钱”“思想是食物”等隐喻，说明我们如何通过具象的经验来理解抽象的概念[10]。

2.3.2. 人类话语与 AI 生成话语的对比

随着 AI 生成内容技术的进步，研究者们开始关注 AI 生成话语与人类话语之间的异同。尽管 AI 生

成的文本在语法和语用方面表现出色, 能够保持高水平的语言一致性和逻辑连贯性, 但在语义深度和情感表达上仍存在明显差异[12]。

人类话语通常包含丰富的语义和情感信息, 语言结构复杂且反映了特定的文化和社会背景。这些特点源于人类在特定语境中的经验、知识和文化认同[13]。例如, 人在描述某一情景时, 通常会结合自身的经历和情感, 从而赋予话语特殊的意义和感染力。

相较之下, 当前的 AI 生成模型主要通过大规模数据的统计学习来产生文本, 尽管能够模拟人类的语言模式, 但对语义理解和情感表达的深度上有所欠缺。这种模式化的结果往往显得机械化, 难以捕捉复杂语境中微妙的情感和文化内涵[14]。此外, AI 生成模型在处理语义歧义和文化特异性方面也存在局限, 这可能导致生成内容在某些情境下不够自然和得体。

综合以上分析, 尽管 AI 生成内容在技术上已经取得显著进步, 但与人类话语相比还存在不少需要改进的地方。未来的研究应继续深入探讨 AI 生成内容的语义和情感特征, 结合认知语言学和话语分析的理论, 为 AI 生成模型的优化提供科学依据。

3. 研究方法

3.1. 研究设计

本研究旨在分析人工智能生成内容中的话语认知特征, 揭示其与人类话语的异同。研究设计采用了混合方法论, 结合了质性和量化方法, 充分利用各自的优势来深入探讨 AI 生成内容在语义、情感和文化背景方面的特点[15]。研究包括三个主要步骤: 数据收集、数据预处理和数据分析。首先, 从多个数据源获取 AI 生成内容和人类话语样本; 然后, 对数据进行清洗和预处理; 最后, 通过质性分析和量化分析揭示话语认知特征。

3.2. 数据收集

3.2.1. 数据来源

本研究的数据来源主要包括三个方面: 1) AI 生成内容, 包括 GPT-3 生成的文本和 DALL-E 生成的图像描述; 2) 人类话语样本, 包括在线社交媒体、新闻报道和文学作品中的文本; 3) 辅助数据, 如语料库和数据库以支持语义分析和情感分析。

AI 生成内容主要通过 API 接口从 OpenAI 的 GPT-3 和 DALL-E 获取。为了确保生成内容的多样性和代表性, 对不同主题和场景下的文本和图像描述进行了采集。人类话语数据则通过网络爬虫技术从公开的社交媒体、新闻网站和电子书平台上收集。这些数据经过筛选和整理, 以确保其内容的广泛性和多样性。

3.2.2. 数据类型

本研究的数据类型主要包括两类: 1) 文本数据, 涵盖 AI 生成的文本和人类话语文本; 2) 图像描述数据, 主要是由 DALL-E 生成的基于文本描述的图像。

文本数据包括自然语言描述、对话、故事和评论等多种形式, 涵盖多个话题和情境。图像描述数据则是根据特定文本内容生成的图像及其文字描述, 这些数据有助于分析 AI 在图像语义理解和生成上的能力。

3.3. 数据分析方法

3.3.1. 质性分析方法

质性分析方法主要用于探索 AI 生成内容中的话语特征及其与人类话语的差异。具体方法包括以下几种:

- 1) 内容分析：通过手动或软件辅助对文本内容进行编码和分类，提取其主题、语义模式和情感表达。内容分析有助于揭示 AI 生成内容在语义结构和情感表达上的共性和差异。
- 2) 语篇分析：采用系统功能语言学(SFL)和批判话语分析(CDA)框架，对话语的功能、结构和社会意义进行深入解读。语篇分析可以帮助理解 AI 生成文本在社会语境中的适用性和连贯性。
- 3) 隐喻分析：运用认知语言学的概念隐喻理论，分析 AI 生成文本中的隐喻使用情况，探讨其在构建抽象意义上的能力。通过比较人类和 AI 生成文本中的隐喻，揭示两者在认知模式上的差异。
- 4) 情感分析：通过自然语言处理技术(NLP)，对文本中的情感进行分类和量化分析。情感分析有助于评估 AI 生成内容在情感表达上的准确性和自然性。

为确保质性分析的可靠性，研究从多名分析师中招募了语言学和计算机科学领域的专家，采用卡帕系数(Kappa Coefficient)测量分析人员之间的一致性，以提高分析结果的可信度。

3.3.2. 量化分析方法

量化分析方法用于统计和比较 AI 生成内容与人类话语在多个维度上的差异。主要方法包括：

- 1) 描述性统计分析：对文本数据的基本特征进行统计描述，包括词频、句长、语法结构等指标。这些描述性统计分析有助于了解 AI 生成内容的语言复杂性和多样性。
- 2) 语义网络分析：利用语义网络构建技术，分析文本中的概念及其关系，揭示 AI 生成内容的语义结构。通过比较人类话语和 AI 生成文本的语义网络，可以发现两者在概念关联和主题分布上的异同。
- 3) 回归分析：采用线性回归和逻辑回归模型，检验文本特征与内容生成质量之间的关系。回归分析有助于识别影响 AI 生成内容质量的关键因素，并为模型优化提供依据。
- 4) 机器学习方法：利用分类和聚类算法，对文本进行自动分类和模式识别。机器学习方法可用于分析和比较不同类型文本在语义和情感特征上的表现，评估 AI 生成模型的性能。

为确保量化分析的有效性，研究采用了大规模数据集，并使用交叉验证(Cross-Validation)技术评估分析模型的稳健性和泛化能力。此外，研究还采用了图像识别和处理技术，对图像描述数据进行辅助分析，以全面揭示文本与图像之间的语义关系。

4. 案例研究

4.1. 样本选择与背景

在本研究中，笔者从多个渠道收集了各种样本，以全面分析人工智能生成内容和人类创作内容之间的认知差异。具体而言，笔者选择了由 OpenAI 的 GPT-3 生成的文本和人类撰写的文本作为分析样本。所选文本涵盖以下几种类型：

新闻报道：选取国际新闻和科技新闻作为主要分析对象，分别从 GPT-3 和知名新闻网站(如《纽约时报》《BBC》)上获取文本样本。每种类型各选取 30 篇，共 60 篇。

社交媒体帖子：主要选取 Twitter 上的社交媒体帖子，通过语料库自动抽样，并由 GPT-3 生成相应的帖子进行对比。每种类型各 30 篇，共 60 篇。

文学作品：选取经典文学作品的段落(如《了不起的盖茨比》《红与黑》)和 GPT-3 生成的文学段落进行对比，每种类型各 30 篇，共 60 篇。

样本总计 180 篇文本，确保多样化和代表性，以便于后续的质性和量化分析。

4.2. 实验设计与过程

本研究采用混合研究方法，包括质性和量化分析，以深入探讨 AI 和人类文本在话语认知特征上的差异。

4.2.1. 数据预处理

数据预处理是本研究的第一步, 旨在为后续分析打下基础。具体步骤如下:

- 1) 文本去噪和标准化: 去除文本的无关部分, 如广告、页脚等。这一步骤使用 Python 中的 NLP(自然语言处理)工具包完成。
- 2) 文本分词和标注: 将文本进行分词处理, 并为每个词语添加词性标注, 以便于后续的语义分析。

4.2.2. 质性分析

质性分析主要通过内容分析、隐喻分析和情感分析三项内容进行。方法如下:

- 1) 内容分析: 利用 NVivo 软件对文本进行编码, 识别主题、语义场域和叙述结构。在新闻报道中, 发现人类文本通常包含背景信息和深入分析, 而 GPT-3 生成的文本更倾向于报道表面信息。
- 2) 隐喻分析: 使用显性隐喻识别工具, 自动提取文本中的隐喻表达。结果显示, GPT-3 生成的文本中每千词隐喻数量为 5~7 个, 而人类文本为 15~20 个, 且人类文本中的隐喻更丰富和复杂。
- 3) 情感分析: 利用情感分析工具(如 TextBlob), 计算文本的情感得分和情感波动。发现 GPT-3 生成的文本情感波动较小, 平均情感得分集中在 0.1 至 0.2 区间, 而人类文本情感波动较大, 得分在-0.5 至 0.5 之间变化。

4.2.3. 量化分析

量化分析则通过描述性统计、语义网络分析和机器学习分类来进行, 具体步骤如下:

- 1) 描述性统计分析: 利用 SPSS 软件, 对预处理后的文本数据进行词频和句长统计。结果显示, GPT-3 生成的新闻文本平均句长为 15 词, 而人类文本为 20 词左右; 词频集中在中高频文本, 而人类文本的词频分布更加均匀。
- 2) 语义网络分析: 使用 Gephi 软件, 构建语义网络并分析文本结构。GPT-3 生成的社交媒体帖子中主要节点集中在少数几个高频词汇, 而人类帖子中节点和边数量显著更多, 网络结构更复杂。
- 3) 机器学习分类: 利用机器学习分类算法(如随机森林和 SVM), 对文本进行分类。分类结果显示, 基于文本特征的分类准确率在人类文本为 85%, 而 AI 生成的文本准确率为 75%, 表明人类文本内在结构的复杂性和多样性。

4.3. 数据展示与初步分析

4.3.1. 质性分析结果

通过质性分析发现:

- 1) 内容与主题: 人类撰写的新闻报道更注重背景和深度分析, 例如一篇关于气候变化的报道不仅涵盖了当前的气候数据, 还讨论了未来趋势和政策影响(来源: BBC)。然而, GPT-3 生成的报道更简短, 缺乏深度, 仅报告了表面的数据和事实。
- 2) 情感表达: 在社交媒体帖子中, 人类文本显示出了广泛的情感波动, 具有高情感强度的词频率显著较高, 如“激动”“愤怒”等。而 GPT-3 生成的帖子更中性, 情感词汇使用频次较低。
- 3) 隐喻使用: 隐喻分析显示, 人类文本中的隐喻表达更为复杂和多样化。例如, 人类创作的文学作品常用隐喻表达情感和背景, 如“黑夜似一张巨大的幕布”来描述夜晚, 而 GPT-3 生成的文本中隐喻使用较少且简单。

4.3.2. 量化分析结果

通过量化分析, 初步结果如下:

- 1) 词频与句长分布: 描述性统计分析显示, GPT-3 生成文本的平均词频较高, 句子长度集中在 15~20

词之间，而人类文本的词频和句子长度分布更加分散，平均词汇多样性指数(TTR)为 0.5，而 GPT-3 生成文本为 0.3。

2) 语义网络结构：语义网络分析表明，人类文本的语义网络结构更为复杂，平均连接度(Average Degree)在人类文本中为 8~10，而在 GPT-3 文本中为 5~6，表明人类文本更能形成复杂的概念连接。

3) 情感得分：情感分析显示，GPT-3 生成的文本情感得分集中在中性区间(平均得分为 0.15)，而人类文本情感得分的变异性较大(平均得分为 0.3)，显示出更强的情感表达能力。

4.3.3. 分类与回归模型

通过分类与回归模型分析，结果显示：

1) 回归分析：线性回归分析表明，文本的逻辑连贯性和词汇多样性是影响新闻文本生成质量的重要因素， R^2 值在 0.65 左右。

2) 机器学习分类：基于机器学习算法的文本分类准确率为 75%~85%，发现情感和互动性是影响社交媒体文本生成质量的关键因素。在人类文本分类中，准确率高达 85%，而 AI 生成文本分类准确率较低，为 75%，表明人类文本更复杂难以复制。

综合质性和量化分析的结果，可以看出 AI 生成内容在某些方面(如逻辑连贯和表面信息报道)接近人类文本，但在情感表达、隐喻使用和语义复杂性上仍存在显著差距。这表明虽然 AI 生成技术在某些领域表现出色，但在话语认知方面仍需进一步优化和提升，以更接近人类的语言表达和认知复杂性。

5. 结果与讨论

5.1. 人工智能生成内容的话语特征

在本部分中，笔者将详细探讨人工智能生成内容的语言复杂度、语义连贯性及情感色彩等话语特征，并通过定量数据加以佐证。

5.1.1. 语言复杂度

在语言复杂度方面，笔者分析了 GPT-3 生成的文本与人类创作的文本在词汇多样性和句子长度等方面差异。词汇多样性通过类型 - 标记比(TTR)进行评估，GPT-3 的 TTR 平均值为 0.4，而人类文本的 TTR 平均值为 0.5，表明人类文本的词汇使用更为丰富多样。

此外，句子长度的统计结果显示，GPT-3 生成的文本平均句长为 15~20 词，而人类文本则为 20~25 词，且分布更为均匀。这说明人类创作者往往使用更长、更复杂的句子结构，而人工智能生成的内容在句子长度上相对较短且规律性更强。

5.1.2. 语义连贯性

为了评估语义连贯性，笔者使用语义网络分析工具对文本进行分词及名词短语提取，构建语义网络并分析其复杂度。通过 Gephi 软件的平均连接度数据显示，人类文本语义网络的平均连接度为 8~10，而 GPT-3 生成的文本为 5~6。这表明，人工智能生成的内容在语义连贯性方面仍有待提升，因为其结构相对简单，语义节点和边的数量较少。

进一步的人工编码分析中，GPT-3 的文本虽能保持局部连贯性，但存在缺乏全局连贯性的情况。例如，在新闻文本样本中，GPT-3 生成的段落往往缺乏对背景信息的深入介绍和逻辑拓展，不如人类新闻文本详细和逻辑紧密。

5.1.3. 情感色彩

通过对情感色彩的分析，笔者使用 TextBlob 工具进行情感分析，分别计算了 AI 生成和人类文本的

情感得分及其波动范围。结果显示, GPT-3 生成的文本情感得分集中在 0.1 至 0.2 之间, 情感波动较小, 相对中性。相反, 人类文本情感波动较大, 其情感得分从 -0.5 到 0.5 不等, 显示出更为丰富且强烈的情感表达。

这一结果也在社交媒体和文学作品样本中得到了验证。比如在人类写作的社交媒体帖子中, 使用的情感词汇更加丰富多样, 而 GPT-3 生成的帖子则显得较为平淡, 缺乏情感深度和多样性。

5.2. 人工智能生成内容的认知分析

在认知分析部分, 笔者探讨了 AI 生成内容对读者认知负荷的影响, 以及其在感知与理解上的特点。

5.2.1. 认知负荷分析

为了评估认知负荷, 笔者分析了不同文本类型对读者的认知负荷感。一方面, 在新闻文本中, 由于 GPT-3 生成内容的平铺直叙和简单结构, 读者在理解过程中往往会感到易读且省力。然而, 缺乏细节和深度会使得读者对信息的全面理解产生不足, 反而增加了二次信息搜寻的认知负荷。

另一方面, 在文学作品中, 人类写作的文本由于丰富的语言表达和复杂的叙述结构, 可能会增加读者的认知负荷, 但这同时也提升了文学欣赏的深度和层次感, 而 GPT-3 生成的文本则因过于简单直接而缺乏文学美感和思考空间。

5.2.2. 感知与理解

在感知与理解上, 读者对 AI 生成的内容普遍感到较为平淡, 缺少情感波动和吸引力。这一现象在情感分析数据中得到了证实。具体而言, 读者对 GPT-3 生成的新闻和社交媒体帖子往往感到缺乏新意和深度, 容易产生疲劳感, 而人类撰写的内容则具备更强的吸引力和持续读者注意力的能力。

在人类编写的文学作品中, 隐喻和象征手法的使用能够激发读者的联想和情感共鸣, 而 AI 生成的文本在这方面则显得较为缺乏。例如, 在经典文学段落中, 隐喻表达如“黑夜似一张巨大的幕布”深刻且富于想象力, 而 GPT-3 生成的句子则相对直白和简单。

5.3. 结果讨论

5.3.1. 与人类生成内容的差异

综合上述分析可见, AI 生成的文本在语言复杂度、语义连贯性和情感表达上与人类文本存在明显差异:

- 1) 语言复杂度: AI 生成的文本语言复杂度较低, 词汇多样性和句子长度均不及人类文本。
- 2) 语义连贯性: AI 文本的语义连贯性较差, 缺乏复杂的语义网络结构, 难以形成与人类文本相一致的连贯性。
- 3) 情感表达: AI 生成的文本情感表达较为中性, 情感波动和多样性远低于人类文本, 影响了文本的吸引力和读者情感共鸣。

5.3.2. 可能的影响因素

上述差异可能受到以下几个关键因素的影响:

- 1) 数据集限制: AI 模型的训练数据集类型和范围直接影响生成内容的质量。若训练数据集缺乏多样性和复杂度, AI 生成的内容自然会有所不足。
- 2) 模型设计: 当前的 AI 生成模型在语言处理能力上仍有局限, 难以理解和生成复杂的逻辑结构和深层次的语义关系。
- 3) 情感理解与表达: AI 在情感理解和表达上的能力有限, 其生成的情感色彩往往较为平淡, 难以达

到人类创作者的水平。

4) 语境与文化背景: 人类创作时会考虑语境和文化背景, 而 AI 生成文本在这方面的考量相对不足, 导致其生成内容的文化深度和语境适应性都较弱。

综合研究结果表明, 人工智能生成内容在话语认知特征上与人类文本存在显著差异, 尽管在某些领域表现优异, 但仍需在语言复杂度、语义连贯性和情感表达等方面进行进一步优化和提升。这不仅为 AI 文本生成技术的改进提供了方向, 也为人工智能在文本生成领域的应用提出了更高的要求。

6. 结论

6.1. 研究的理论贡献与实际意义

6.1.1. 理论贡献

首先, 本研究为人工智能生成文本的话语分析提供了系统的理论框架, 丰富了话语分析领域的传统理论。从语言复杂度、语义连贯性和情感色彩等多个维度对 AI 生成内容进行系统分析, 揭示了 AI 生成文本与人类文本在多方面的异同。这些发现不仅验证了某些现有理论, 也为进一步探索 AI 在话语生成中的认知特征提供了实证数据。

其次, 研究结果指出了当前人工智能生成技术在语言处理上的不足, 为改进和优化 AI 生成内容提供了理论依据。例如, 针对 AI 生成内容词汇多样性较低的问题, 可通过丰富训练数据集和引入高级语言模型来提高生成内容的语言复杂度。针对语义连贯性不足的问题, 则需在 AI 模型中引入更多层次的语义网络分析, 优化智能生成算法。

6.1.2. 实际意义

在实际应用方面, 本研究结果对 AI 文本生成技术在新闻、文学创作、社交媒体和智能助手等领域的应用具有重要指导意义。在新闻领域, 了解 AI 生成文本的语言特征和认知负荷, 有助于新闻从业者在使用 AI 写作工具时优化文本质量, 提高信息传达的清晰度和连贯性。在文学创作领域, 研究结果可以帮助文学作者和出版商探索 AI 辅助创作的新可能性, 同时也规避由于 AI 语言表达不足可能带来的文化深度和情感共鸣缺失的问题。

在社交媒体和智能助手应用中, 理解 AI 生成内容的情感表达特点, 可以帮助开发者增强 AI 在情感交互和用户体验方面的能力, 提高用户对 AI 生成内容的接受度和喜好。

6.2. 研究局限与未来研究方向

6.2.1. 研究局限

本研究存在以下几个局限:

首先, 由于数据样本有限, 研究结果可能无法全面代表所有 AI 生成文本的特征。不同领域和不同文本类型的特点差异较大, 需要更为广泛和多样的数据样本进行进一步验证和补充。

其次, 本研究主要基于定量的语言学和认知分析方法, 而未能深入探讨 AI 生成文本在具体应用场景中的实际效果和用户体验。这些方面的定性研究需要进一步开展, 以全面理解 AI 文本生成的特点和应用影响。

最后, 本文未能充分考虑不同 AI 模型之间的差异。不同 AI 模型在语言处理和生成上的表现差异较大, 研究应进一步细化, 分析不同模型在生成内容上的特定话语特征及其影响因素。

6.2.2. 未来研究方向

未来的研究可以在以下几个方面进一步深入:

一是扩大数据样本范围，涵盖更多领域和类型的 AI 生成文本，从而获得更具代表性的研究结果。例如，可以引入更多来自新闻、娱乐、科技等不同领域的文本样本，进行更加细致的分类和对比分析。

二是结合定性研究方法，深入探讨 AI 生成内容在特定应用场景中的实际效果和用户体验。通过访谈、问卷调查等方法，获取用户对 AI 生成内容的主观评价和反馈，以补充定量分析的不足。

三是研究不同 AI 模型在文本生成上的差异及其影响因素，探索优化 AI 文本生成的具体路径。例如，可以比较 GPT-4 与其他高级别语音生成模型(如 BERT、T5 等)在不同文本生成任务中的表现，并探讨不同模型结构和训练方法对生成文本特征的影响(Murphy, 2012)。

通过以上途径，未来的研究将能够更加全面和深入地理解和优化人工智能生成文本的能力，从而推动 AI 文本生成技术在实际应用中的广泛和深入应用。

参考文献

- [1] Brumby, D. (2004) Artificial Intelligence: A Modern Approach. *Telecommunication Journal of Australia*, **54**, 63-64.
- [2] Nilsson, N.J. (1996) Artificial Intelligence: A Modern Approach. *Artificial Intelligence*, **82**, 369-380. [https://doi.org/10.1016/0004-3702\(96\)00007-0](https://doi.org/10.1016/0004-3702(96)00007-0)
- [3] Cao, C., Liu, F., Tan, H., Song, D., Shu, W., Li, W., et al. (2018) Deep Learning and Its Applications in Biomedicine. *Genomics, Proteomics & Bioinformatics*, **16**, 17-32. <https://doi.org/10.1016/j.gpb.2017.07.003>
- [4] Seo, J., Moon, H., Lee, C., Eo, S., Park, C., Kim, J., et al. (2022) Plain Template Insertion: Korean-Prompt-Based Engineering for Few-Shot Learners. *IEEE Access*, **10**, 107587-107597. <https://doi.org/10.1109/access.2022.3213027>
- [5] Sun, G., Liang, W., Dong, J., Li, J., Ding, Z. and Cong, Y. (2024) Create Your World: Lifelong Text-to-Image Diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **46**, 6454-6470. <https://doi.org/10.1109/tpami.2024.3382753>
- [6] Child, G.S. (2019) Book Review: James Paul Gee, an Introduction to Discourse Analysis: Theory and Method (Fourth Edition). *Qualitative Research*, **19**, 236-238. <https://doi.org/10.1177/1468794116683992>
- [7] Halliday, M.A.K. and Hasan, R. (1976) Cohesion in English. Longman.
- [8] Halliday, M.A.K. (2004) An Introduction to Functional Grammar. Hodder Education Bookpoint Limited.
- [9] Fairclough, N. and Nekvapil, J. (2006) RECENZE—Analysing Discourse. Textual Analysis for Social Research. *Socio-logicicky Casopis*, **65**, 435-438.
- [10] Lakoff, G. and Johnson, M. (2003) Metaphors We Live By. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226470993.001.0001>
- [11] The Linguistic Society of Korea (1982) Linguistics in the Morning Calm. Hanshin Publishing Company.
- [12] Bender, E.M., Gebru, T., McMillan-Major, A. and Shmitchell, S. (2021) On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 3-10 March 2021, 610-623. <https://doi.org/10.1145/3442188.3445922>
- [13] Tannen, D. (1993) Framing in Discourse. Oxford University Press.
- [14] Goldfarb, D. (2020) Rebooting AI: Building Artificial Intelligence We Can Trust. *Computing Reviews*, **61**, 289.
- [15] Creswell, J.W. and Plano Clark, V.L. (2007) Designing and Conducting Mixed Methods Research. Sage Publications.