

# 人工智能助力阅读体验：沉浸式智慧阅读服务模式创新<sup>\*</sup>

梁建春

**摘要** 论文针对当今图书馆在阅读体验、个性化推荐和智能问答等方面的不足,提出一种以AI为核心的多模态“沉浸式智慧阅读体验框架”,该框架引入富媒体概念,能较好弥补阅读双编码理论缺乏“文本和图像之外的其它事物”认知可能性的不足。结合多模态输入输出、用户画像和知识图谱生成等技术,构建AI多模态沉浸式智慧阅读服务模型。“AI+阅读”的融合变革,不仅较好弥补并完善了双编码理论在实践应用方面的不足,而且能够提升阅读体验,推动人工智能与阅读领域的交叉融合,促进阅读研究与实践的创新。

**关键词** 智慧图书馆;人工智能;双线理论;沉浸式阅读;ChatGPT;机器学习

**分类号** G250.7

**DOI** 10.16810/j.cnki.1672-514X.2023.12.009

## AI Empowering Reading Experience: Innovation in Immersive Intelligent Reading Service Model

Liang Jianchun

**Abstract** This paper first addresses the shortcomings of some intelligent libraries in aspects such as reading experience, personalized recommendations, and intelligent Q&A. It proposes an AI-centered multi-modal “Immersive Intelligent Reading Experience Framework”, which introduces the concept of rich media and helps to bridge the gap in the dual-coding theory’s lack of cognitive possibilities “beyond text and images”. Subsequently, by incorporating multi-modal input-output, user profiles, and knowledge graph generation technologies, an AI multi-modal immersive intelligent reading service model is constructed. Through the integration of “AI + Reading” transformation, not only are the shortcomings of the dual-coding theory in practical applications addressed, but also the reading experience is enhanced. This promotes the cross-fusion of artificial intelligence and the field of reading, facilitating innovation in reading research and practice.

**Keywords** Smart library. Artificial intelligence. Dual-coding theory. Immersive reading. ChatGPT. Machine learning.

### 0 引言

双编码理论 (Dual Coding Theory) 是 Paivio 在 1971 年首次提出的一种认知理论,认为视觉和语言信息都有助于阅读学习和记忆。人脑中的不同通道会对这些信息进行不同的处理,并为每个通道中处理的信息创建单独的心理代码。与仅以一种方式对信息进行编码相比,以两种不同方式进行编码的能力增加了记住该信息的机会。例如:当我们阅读一段描述沙滩的文字时,可以同时想象出沙滩的图像,这样可以帮助读者更好地理解 and 记忆这个场景。双编

码理论的重要性在于可以更好地解释人类的阅读和学习过程。但该理论的局限性在于未考虑到认知可能由文本和图像以外的事物促成,忽略了人们对声音、触觉和味觉等其他感官的输入认知,而且也未能充分考虑文字识别和语言理解之间的相互影响,因此并不能完全解释阅读能力提升原因的复杂性。

要解决双编码理论的局限性,需要考虑包括视觉和语言在内的一切认知形式,例如声音、感觉和情感等,同时还需要探索多种方式进行编码,例如音频、图形、视频等,这将有助于更全面地理解记忆和学习的过程。随着 AI 技术

<sup>\*</sup>本文系2023年广西师范大学图书馆馆内规划项目“信息化时代下高校智慧图书馆服务创新实践研究——以广西师范大学为例”(项目编号:TSG202315)阶段性研究成果之一。

在全球范围多领域的广泛应用，可以利用自然语言处理、计算机视觉和语音识别等关键技术构建具有多模态特征的“AI+ 阅读”学习和记忆模式，从而对双编码进行扩展，形成多编码的阅读认知，增强和提升阅读记忆及阅读效率。

## 1 国内外图书馆阅读服务现状

以创新的教育和技术服务著称的美国芝加哥哈罗德·华盛顿图书馆，提供各种数字资源、创客空间及计算机和编程课程，采用基于 AI 的推荐系统可以根据用户的阅读历史和兴趣推荐相关内容，还能通过虚拟现实和增强现实技术为读者提供沉浸式阅读和学习体验。但在智能问答系统、语音识别等方面的应用相对较弱。美国费城图书馆设有创客空间，提供 3D 打印、虚拟现实等服务，在线资源和远程访问服务能够极大方便读者获取知识信息。然而它没有充分利用人工智能技术改进阅读推荐和用户体验。

国内一些大型图书馆引入了自然语言处理技术，通过对读者查询语言的分析，智能化地推荐相关图书和资料。这种应用的优势在于能够帮助读者快速、精准地找到所需材料，提升阅读效率和体验。不足之处是在处理语义不明确或者模糊查询时会有推荐不准确的情况。另有一些图书馆引入了机器人服务，如可移动的阅读推广机器人、自动导航机器人等。这种应用的优势在于能够提供 24 小时服务，为读者带来具有时效性的阅读服务。不足之处是机器人在服务过程中无法提供情感交流和个性化服务，可能会影响读者的互动体验。

Walton<sup>[1]</sup> 在研究智能图书馆是否会改变阅读和学习模式的过程中认为，智能阅读可以提高信息获取、处理和利用的效率，有助于提高用户的阅读体验和学习成果。国内学者张瑜<sup>[2]</sup> 将智慧阅读视为一种基于互联网的新型阅读方式，强调了个性化推荐、用户参与等特点。孙艳<sup>[3]</sup> 在探讨公共图书馆智慧阅读服务平台构建的同时，将智慧阅读定义为一种基于移动互联网和大数据的阅读模式，重点关注了阅读服务的便捷性、个性化等方面。尽管国内外图书馆在智慧阅读平台的数字资源建设、算法推荐、创新服务上取得了一系列的成果，但仍需要进一步加强人机交互体验、分析读者阅读情感、优化推荐算法、

扩展智慧阅读的功能范畴和生成环境<sup>[4-7]</sup>。

综上所述，国内外大部分图书馆使用了初步的 AI 技术，主要是智能推荐，智能导航，机器人简单问答（问题和回答都是在事先设置好的范畴内），少数图书馆用虚拟现实技术来完成一些相关业务展示。但是对于阅读，尤其是 AI 沉浸式阅读领域，很少做过详细的体系框架和模型扩展研究。ChatGPT4.0 的正式发布和利用 AI 核心衍生的一系列文本、图形、图像和视频处理产品的实践应用，是人工智能领域的转折性的突破，为图书馆打造更加丰富的阅读体验提供了可行性。因此，本文在构建 AI 沉浸阅读框架基础上，把现有的 AI 关键技术整合在一个模型之中，采取应用场景插件式模块化组合，可以根据环境和经费选择或添加场景插件，构建多模态沉浸式智慧阅读模型。

## 2 多模态沉浸式智慧阅读构想

沉浸式智慧阅读是指将虚拟现实、增强现实、自然语言处理、机器学习等技术与传统阅读相结合，创造出一种更加丰富、生动、互动的阅读体验。在实践应用上具备三个优势：（1）通过语音合成技术将文本转化为语音，并加入情感色彩和语音音调的调节，引起读者情感共鸣，深入理解作者意图；（2）通过增强现实和虚拟现实技术，将文本呈现在更加真实、立体的场景中，增强阅读的体验感和可视化效果；（3）根据读者的个性化需求和兴趣，提供更加智能化的阅读体验，例如推荐相似主题、翻译、注释、词汇扩展等。

在 AI 中，多模态（multimodality）通常以多种传感器、算法和模型来实现。基于 AI 的多模态属性，可以帮助读者更好地理解 and 掌握阅读信息。读者可以通过图像、视频、语音、语言和其他传感器来获取不同类型的信息，并使用 AI 算法和模型将这些信息整合在一起。这样，读者就可以通过多个感官通道来处理和理解信息，以获得更好的阅读体验和学习效果。

以“AI 多模态+沉浸式阅读”的认知模式分析和讨论为出发点，笔者构想了多模态沉浸式智慧阅读的主要思路。（1）图像识别和语义分析：通过图像识别技术，将图书中的图片和文字进行分离和识别，再通过自然语言处理和语义分析等深度学习，将文字内容和图片内容进行关联和整

合, 提高阅读体验。(2) 视觉呈现优化: 使用“AI+图像处理”技术对阅读界面的优化, 提高文字的可读性和视觉效果, 如调整字体大小、行距、背景色、文字颜色等。(3) 语音识别和自然语言处理: 使用语音识别技术, 将文字内容转化为语音, 也可以通过自然语言处理, 将语音内容转化为文字, 实现文字内容和语音内容的双向转化。(4) 虚拟现实技术: 利用虚拟现实技术, 将文字和图像等信息以更加直观的方式呈现给读者, 提高双编码理论在阅读中的应用效果。

在编码实现方面, 需要根据具体的应用场景和需求, 选取合适的编程语言和开发平台, 如 Python、TensorFlow、PyTorch 等。在开发过程中, 需要考虑数据的采集、预处理和模型的训练和测试等问题, 同时需要注意模型的效率和精度。

### 3 以 ChatGPT-4 为内核的 AI 沉浸式阅读体验框架

ChatGPT 是 OpenAI 开发的一款基于深度学习的大型语言模型聊天机器人, 目前发布的 ChatGPT-4, 不仅能够与人类进行自然、流畅和有趣的对话, 而且还具有生成创意内容、提供信息、执行任务等多种功能, 是目前最先进的聊天机器人之一, 展示了人工智能在自然语言处理领域的巨大潜力。

ChatGPT-4 的衍生产品主要有两类: 一类是利用其文本生成能力, 结合图形图像处理或视频处理技术, 实现多媒体内容的自动创作; 另一类是利用其对话交互能力, 结合语音识别或语音合成技术, 实现自然语言理解和响应的智能助理。例如, Whisper 是一个基于 ChatGPT 和 DALL·E 的多媒体内容创作平台, 用户可以通过简单的文字指令, 生成各种类型和风格的图片、视频、音乐等。用户可以输入“画一只穿着西装打领带的猫”, Whisper 就会生成相应的图片。而 ChatGPT Plus 是一款融合 ChatGPT 和 InstructGPT 的高级聊天机器人服务, 用户可以通过网页或手机应用与之对话, 并获取各种信息和服务。用户可以输入“请帮我订明天上午 11 点从哈尔滨到广州的高铁票”, ChatGPT Plus 就会为用户完成订票操作, 并提供相关信息。ChatGPT Voice 则是一个基于 ChatGPT 和语音技术的智能助理应用, 用户可以通过语音与之交流, 并获取各

种信息和服务。用户说“请告诉我今天北京的天气情况”, ChatGPT Voice 就会回答“今天北京多云转晴, 最高温度 22 度, 最低温度 10 度”。

将 ChatGPT-4 内核功能及相关衍生产品整合融入到 AI 多模态沉浸式智慧阅读模式构想之中, 可以促进图书馆和人工智能(AI)共同发展, 相得益彰。一方面, AI 可以为沉浸式智慧阅读平台提供强大的自然语言处理能力, 使得读者和读者之间、读者和平台之间进行自然、流畅的交互。另一方面, 阅读平台可以为 AI 提供更加丰富、多样化的文本语料库, 不仅包括传统的书籍、期刊等, 还包括虚拟现实、增强现实等环境下生成的语言数据。AI 可以在更多的文本数据上进行训练和优化, 提高其自然语言生成能力和准确性。

基于以上思路, 本文提出了一种基于 ChatGPT-4 技术和功能内核的多模态沉浸式智慧阅读体验框架 (Multimodal Immersive Intelligent Reading Experience Framework, MIREF), 如图 1 所示。

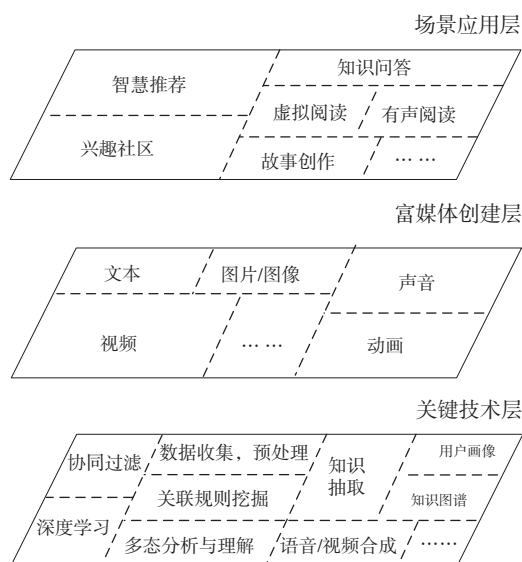


图 1 多模态沉浸式智慧阅读体验框架

MIREF 目的在于利用 AI 的多模态生成属性, 对传统的图书馆阅读体系进行优化和完善, 通过多元化阅读为读者提供个性化、互动、有吸引力的阅读体验, 以满足不同类型的读者和场景需求。

一方面, MIREF 整合多种技术和数据源, 包括情感认知模型、自然语言处理和深度学习等, 以实现读者行为和环境的综合分析和理解。与传统的阅读研究相比, MIREF 能克服以



往研究中对于单一数据源的依赖，较好地解决了阅读体验 and 用户满意度的问题。

另一方面，MIREF 通过整合不同模态的数据，如读者交互数据、生物反应数据和环境数据，实现了对多模态数据的处理和分析。这种

多模态数据处理的应用为阅读体验的个性化和智能化提供了新的视角和解决方案，为探索多模态数据处理在阅读领域的应用开辟了新的途径。图 1 中的不同场景实现所需要的关键技术描述见表 1 所示。

表 1 MIREF 中各场景实现的关键技术列表

| 主题场景 | 场景应用   | 关键技术支持  |
|------|--|---|
| 智能推荐 | 聊天机器人作为阅读助手，帮助读者与他人交流、讨论阅读内容、发现新书籍并分享阅读体验。                                       | ① 自然语言处理：理解和生成自然语言，分析读者输入并生成恰当回应。<br>② 读者画像建立：收集和分析用户数据，以理解用户的兴趣和偏好。<br>③ 情感分析：分析读者情感，提供更个性化的回应和建议。<br>④ 协同过滤和深度学习：利用算法和用户数据为读者提供喜爱的资料。   |
| 兴趣社区 | 聊天机器人可以协助创建和维护兴趣社区，促进读者就特定主题展开讨论和分享经验。   | ① 话题建模：分析讨论内容，提取关键主题，便于读者快速了解重点。<br>② 社区管理：监测和处理不良内容。   |
| 知识问答 | 人工智能创建互动内容，例如问答游戏、填空题和知识竞赛，让读者在轻松愉快的氛围中学习和探索书籍。                                  | ① 文本摘要：从大量内容中提取关键信息，生成简洁的摘要。<br>② 知识图谱：利用知识图谱的关联性，生成有趣的问题和挑战。<br>③ 对话管理：引导和维护有意义的多轮对话。<br>④ 语义关联：识别相关的概念和观点，促进有深度的讨论。   |
| 虚拟阅读 | 增强和虚拟现实设备将读者置身于一个虚拟的环境中，与书籍内容进行更深入地互动，从而体验沉浸式阅读的快乐。                              | ① 3D 建模与渲染：创建和渲染逼真的虚拟环境。<br>② 跟踪与交互：实现头部、手部和眼球跟踪，用户在虚拟空间进行自然交互。<br>③ 文本识别与解析：将书籍内容转化为适合 AR/VR 环境的格式。<br>④ 语音识别与合成：为用户提供语音输入输出。<br>⑤ 聊天机器人技术：AI 在虚拟环境中回答读者问题。  |
| 有声阅读 | 通过音频、视频等多种形式为读者提供阅读体验。例如，通过有声书，让读者在无法阅读的情况下聆听书籍内容。                               | ① 语音识别：将用户的语音指令转化为文本。<br>② 语音合成：将书籍内容转化为音频，以便聆听。<br>③ 文本处理与解析：分析和处理文本内容，使其适应多媒体形式。<br>④ 聊天机器人技术：让 AI 提供内容解释、回答问题和提供附加信息。  |
| 故事创作 | 聊天机器人可生成故事灵感；进行角色设定、情节安排等。动态生成故事，能够实时响应读者输入，根据不同的选择生成不同的故事发展和结局，为读者提供丰富的沉浸式阅读体验。 | ① 自然语言理解：让 AI 理解读者的输入，包括情感和选择，以便在故事中进行相应调整。<br>② 自然语言生成：根据读者的互动和选择，生成多样化的故事走向和结局。<br>③ 上下文感知与连贯性：确保能够跟踪故事中的角色、情节和事件，以保持故事的连贯性和一致性。<br>④ 逻辑推理与情节一致性：根据读者输入，进行逻辑推理和情节安排，生成合理的故事走向和结局。<br>⑤ 多样性与创意：利用深度学习和神经网络，为读者提供多样化故事体验。 |

由表 1 可知，MIREF 将以 ChatGPT、百度文心一言为代表的 AI 聊天机器人内核技术融入图书馆沉浸式智慧阅读体系之中，可以形成灵活、丰富的沉浸式阅读场景。

首先，多模态差异化内容生成可满足读者阅读内容智能推荐需求。通过对历史阅读记录、搜索历史、社交媒体行为等多维度大数据分析，了解读者兴趣爱好、阅读偏好及对不同元素的反应，从而生成更贴近读者需求的情节。

其次，读者在一定程度上自由定制角色、情节和故事背景等元素，实现更高程度的沉浸感和自我参与感。例如，读者可以自定义故事中的主人公、配角、场景等，甚至可以选择不同的结局和发展方向，以获得更加符合自己兴趣和偏好的阅读体验。另外，利用 AI 技术，还可以为读者提供实时、准确的问题解答，帮助读者深入理解所阅读的内容。

再次，富媒体是一种通过结合多元媒体元素，如图像、音频、视频、动画等，来呈现更

加生动、多维和交互式内容的数字媒体形式，有助于增强读者的视觉和感性冲击力，为读者提供更为丰富的沉浸式体验。

从次，阅读对象不再是传统的纸质书籍，沉浸式智慧阅读服务可在多种媒介终端设备上使用，不受空间和时间限制，不同的读者需求和场景，都可以得到定制体验。

最后，建立共同兴趣的阅读社区。读者可以在论坛或群组中发帖、回帖，通过分享自己的阅读体验和见解与其他人互动；也可以通过在线聊天的方式交流阅读感受和心得；还可以通过虚拟角色的对话和扮演，参与到故事创作和演绎中去，增强沉浸式阅读的体验和趣味性。通过社区互动，读者可以更好地了解 and 探讨各种阅读话题和文化现象，形成个性化的阅读品味和态度，推动阅读文化的繁荣和发展。

4 AI 多模态沉浸式智慧阅读模型构建

在 MIREF 框架指导下，本文构建一种

基于多模态数据处理、自然语言处理和深度学习的 AI 多模态沉浸式智慧阅读模型 (AI multimodal immersive intelligent reading model, AI\_MIIRM), 以提高阅读效率和用户满意度。AI\_MIIRM 包括用户界面与交互设计、数据收

集与预处理、文本分析与理解、知识抽取与结构化、内容推荐、阅读辅助、学习与评估等功能模块, 模块之间相互协作, 以逻辑调用的方式为读者提供个多模态、个性化服务, 如图 2 所示, 各模块功能与相互逻辑关系如下。

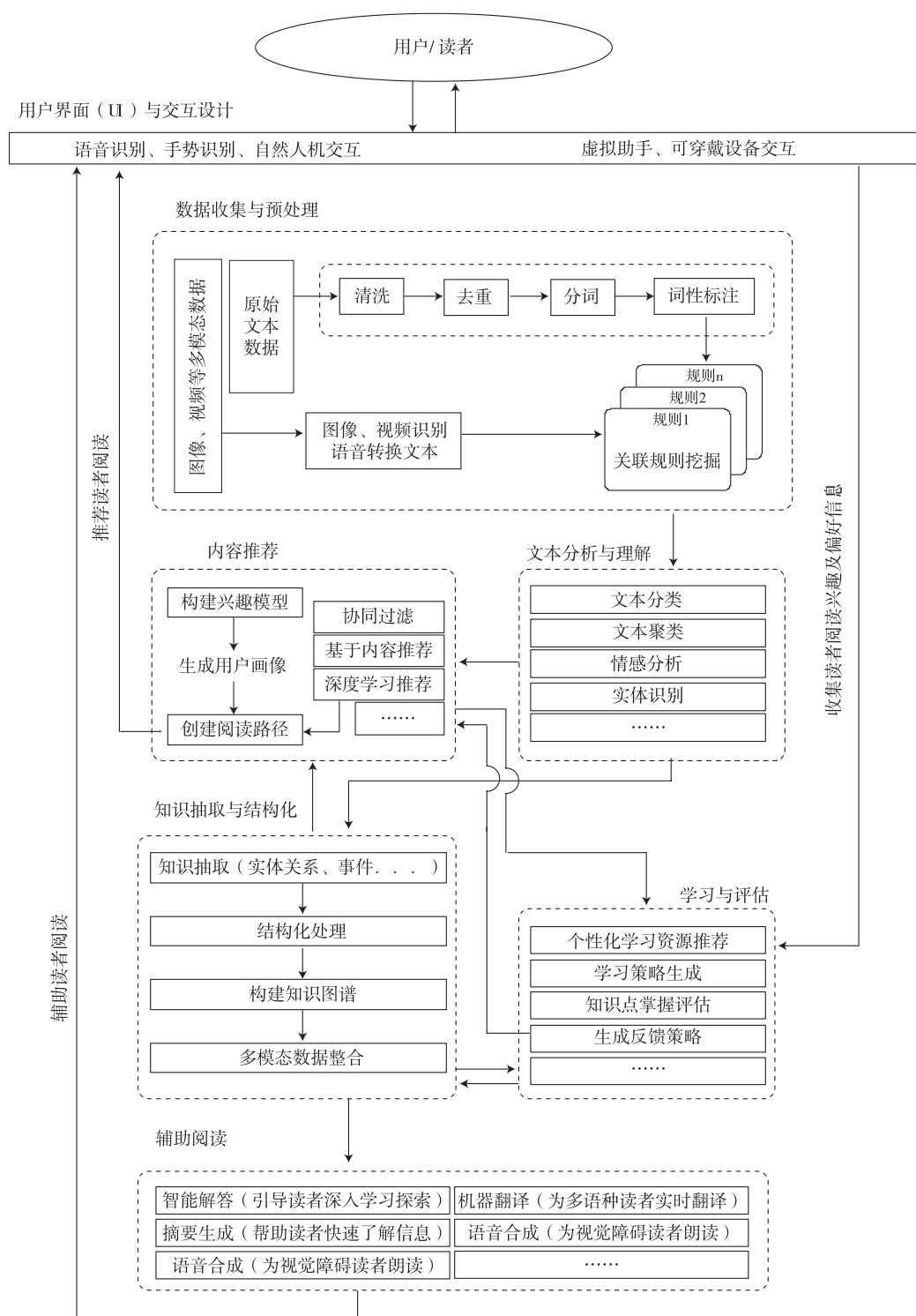


图 2 AI 多模态沉浸式智慧阅读模型 (AI\_MIIRM)

#### 4.1 数据收集与预处理

这是整个沉浸式智慧阅读模型的基础。一方面，它负责收集各种类型文本数据，如书籍、文章、论文等，对数据进行预处理，包括清洗、去重、分词、词性标注等；另一方面，还能收集多模态数据，如图片、音频和视频等。通过这种方式，沉浸式智慧阅读模型可以更全面地理解和分析不同类型的信息。数据收集之后，则需要对文本，以及一些其它多模态数据进行相应的预处理，如图像识别、语音转换为文本等。

#### 4.2 文本分析与理解

接收来自数据收集与预处理模块的文本数据，利用自然语言处理技术进行句法分析、语义分析，主要是利用命名实体识别、关键词抽取、情感分析等技术进行文本生成、理解和推理等。对于图像、视频等多模态数据，可以使用图像、视频识别等技术来识别其中的实体、场景和关系，从而丰富文本分析结果。完成分析后将结果传递给知识抽取与结构化模块和内容推荐模块。

#### 4.3 知识抽取与结构化

从文本分析与理解模块接收分析结果，使用实体关系抽取、事件抽取等技术进行知识抽取和结构化处理。将抽取的知识构建成知识图谱，便于可视化展示和检索，从而为阅读辅助、学习与评估等提供支持。在构建知识图谱时，可以将多模态数据的分析结果整合到知识图谱中。例如，可以将图像和视频识别出的实体和关系添加到知识图谱中，从而使知识图谱更加丰富和多元化。

#### 4.4 内容推荐

基于文本分析与理解的结果，分析读者行为数据，构建读者兴趣模型，为读者生成详细的个人阅读特征描述——用户画像。在构建用户画像时，我们可以综合分析用户在多种类型数据上的行为和偏好，例如可以分析用户在观看视频、阅读图书和听音频时的兴趣和偏好，通过对多模态数据的分析，用户画像将更加精准和全面，从而提高推荐的准确性和个性化程度。随后，基于协同过滤、内容推荐、深度学习推荐等技术自动生成匹配用户画像的阅读路径，进行个性化内容推荐，推荐结果将呈现给读者，并同用户界面与交互设计进行交互。

#### 4.5 阅读辅助

利用知识抽取与结构化的知识图谱及其他相关技术为读者提供多语种辅助阅读。将知识图谱中的实体和关系以可视化的形式呈现给读者，帮助读者更好地理解概念之间的关系。同用户界面与交互设计协同工作，为读者提供友好的交互体验。如“智能解答”可以通过知识图谱帮助 AI 更有效地组织和检索知识，理解不同领域之间的联系，更好地解答读者的问题，引导读者进行深入的学习和探索。“摘要和概述生成”帮助读者快速了解图书或文献的主要信息。“机器翻译”为多语种读者提供实时翻译、语言辅导和学习建议等服务。此外，语音合成技术也可以用于为有视觉障碍的用户提供朗读服务。

#### 4.6 学习与评估

利用知识抽取与结构化产生的知识图谱、用户行为数据及推荐的结果，为用户提供个性化的学习资源和策略。在对读者知识点掌握情况进行评估时，要将评估结果反馈给读者，以便读者调整学习策略。同时，引入领域专家对模型进行指导和评估，确保生成的内容质量和逻辑性。

#### 4.7 用户界面与交互设计

这是读者与系统进行交互的桥梁。它负责呈现推荐内容、提供阅读辅助功能，并收集读者行为数据。它与内容推荐、阅读辅助等紧密协作，通过友好、易用的用户界面，使用语音识别、手势识别等技术实现自然人机交互，为读者提供良好的用户体验。

### 5 AI\_MIIRM 机器学习过程实现

在 AI\_MIIRM 中，机器学习扮演着至关重要的角色。AI\_MIIRM 通过算法让计算机从大量的阅读数据中学习，自动识别出数据中的模式和规律，并根据这些模式和规律做出有意义的阅读和推荐决策。AI\_MIIRM 的机器学习流程包括数据采集、数据预处理、特征选择、模型训练、模型评价和模型调优等多个环节。

#### 5.1 数据采集

数据采集是整个机器学习流程的基础。除了文本、图像、音频、视频等不同类型的数



者的阅读体验。

对于读者的交互数据,可以通过记录读者的点击、滚动、注视等行为,来了解读者对于不同内容的关注度和兴趣点,进而调整阅读界面和内容,提升阅读体验。生物反应数据方面,则可以通过记录读者的心率、脑波等生理反应数据,来了解读者的阅读情绪和认知负荷,从而调整阅读体验和-content,进一步提升阅读质量。在环境数据方面,需要记录光线、温度、湿度等因素,以了解环境对于阅读体验的影响,从而在设计阅读界面和内容时进行优化。不同类型的数据可以使用相应的库和方法进行读取和处理,例如使用 pandas 库的 read\_csv 函数读取 CSV 格式的文本数据,使用 Python 的 OpenCV 库或 Pillow 库进行读取和处理图像数据,使用 Python 的 Librosa 库或 Pydub 库进行读取和处理音频数据,使用 Python 的 OpenCV 库进行读取和处理视频数据。数据采集的目的是获取足够的样本数据,并进行预处理和特征提取。

## 5.2 数据预处理

在数据预处理环节中,可以处理各种类型的富媒体数据,这些数据在不同场景下有着广泛的应用。例如:在图像识别方面,可以处理图像数据,用于人脸识别、目标检测、图像分割等领域。在文本处理方面,可以处理文本数据,用于情感分析、自然语言处理、关键词提取等领域。在音频处理方面,可以处理音频数据,用于语音识别、情感分析、音乐分类等领域。在视频处理方面,可以处理视频数据,用于行为识别、动作分析、人体姿态估计等领域。除此之外,还可以处理一些特定场景下特殊类型的数据。例如,处理读者的交互数据、生物反应数据和环境数据来优化阅读体验。

## 5.3 特征选择

AI\_MIIRM 需要考虑对于阅读体验具有最大影响力的特征。例如:对于电子书阅读,可能需要选择与页面布局、字体、字号、颜色等相关的特征,以及与读者交互相关的特征。例如:翻页时间、阅读速度、书签、高亮等。同时还需要考虑读者的个性化需求,如读模式、调整字号、设置语言等。这些特征都可以通过数据预处理和特征选择来提取并输入机器学习模型中进行训练和预测。可以使用 SelectKBest

函数选择互信息排名前 k 个特征作为输入数据,以保留与沉浸式阅读相关的最重要特征。

当然,也可以选择交互数据、生物反应数据及环境数据中最具有代表性的特征,以便进行后续的模型训练。例如,可以使用点击、滚动、注视等读者交互数据,结合机器学习算法进行特征选择,以选择对于沉浸式阅读相关性最高的交互特征。读者的心率、脑波等生物反应数据也可以作为特征选择的依据,例如,我们可以使用基于心率和脑电图的情感分析算法,对读者的生物反应数据进行分析 and 特征提取,进而结合机器学习算法进行特征选择,以选择对于沉浸式阅读相关性最高的生物反应特征。

## 5.4 模型训练

AI\_MIIRM 使用一些机器学习算法对经过特征选择的数据进行训练。对于沉浸式阅读体验的需求,可以使用随机森林分类器对训练集进行训练。随机森林分类器可以针对多模态数据的特点,具有较好的分类和预测能力,预测读者的兴趣点和情感状态,并根据这些预测结果进行相应的调整和优化。例如:可以使用机器学习算法对读者交互数据和生物反应数据进行训练,以预测读者对于不同内容的关注度和兴趣点,并结合阅读界面和内容进行相应的调整和优化,以提升阅读体验和质量。

## 5.5 模型评价

为对模型进行改进和调优,可以使用三个指标(准确率、召回率、F1 分数)来评估训练好的模型的性能。准确率是分类器正确分类的样本数占总样本数的比例,召回率是指分类器正确识别出的正样本数占总正样本数的比例,F1 分数是准确率和召回率的调和平均数,用于衡量分类器的综合性能。

AI\_MIIRM 为了计算这些指标,需要将模型用于一个包含读者交互数据、生物反应数据和环境数据的独立测试集,并与测试集中的真实标签进行比较。通过对这些指标的评估,可以得出一个关于模型性能在不同情境下的整体评估。常用的机器学习工具库例,如 Scikit-learn 提供了很多函数和工具来计算这些指标,可以方便地使用这些工具来评价模型的性能。在沉浸式阅读环境下,模型评价可以帮助我们了解模型在不同的交互、生理和环境条件下的表现,

为改进和调优模型提供更具有针对性的方向。

## 5.6 模型调优

为了使模型能更好地适应读者阅读行为和偏好，在调优过程中，可用读者行为数据和生物反应数据来指导模型参数的调整。通过记录读者的点击、滚动和注视等行为，以及生物反应数据如心率、脑波等数据，了解读者的阅读兴趣和情绪状态，从而根据不同读者的需求和偏好进行模型参数的调整。另一方面，需要注意过拟合问题。在沉浸式阅读中，可能存在数据量不足或者数据质量较差等问题，这可能导致模型在训练集上表现良好但在测试集上表现不佳。为了避免过拟合，可以采用正则化方法，或者增加数据量、降低模型复杂度等来调整模型。

通过上述 AI\_MIIRM 机器学习环节，运用大量的阅读数据和相关信息，可构建一个能够自动识别出数据中的模式和规律，并根据这些模式和规律做出有意义的阅读和推荐决策的机器学习模型。同时，我们也可以结合读者的交互数据、生物反应数据和环境数据，进一步提高模型的性能和适应性，使得读者能够更加舒适、高效、深入地进行阅读。

## 6 结语

构建多模态沉浸式智慧阅读模型在实践中具有广泛的应用价值，不仅可以改变我们的阅读方式，提高阅读效率，且有助于推动阅读相关技术的发展。

通过整合多种技术和数据源，MIIREF 框架

能够结合读者的情感、认知和环境等多个因素，为其提供更加沉浸、高效的阅读环境，从而提升阅读体验。不仅如此，MIIREF 将为阅读应用领域带来新的发展机遇。其应用可以扩展到图书馆、在线阅读平台、教育培训等领域，为用户提供更加智能、个性化的阅读服务。

MIIREF 将人工智能技术与阅读领域进行有机融合，通过深度学习、自然语言处理、多模态数据处理等技术的应用，可以实现对读者行为、情感、环境等多维度数据的分析和理解，从而提供个性化、智能化的阅读服务。这将推动人工智能在阅读领域的应用和发展，并拓展人工智能在实际场景中的应用范围。

多模态沉浸式智慧阅读框架的构建为阅读研究和实践带来新的思路和方法，不仅可以推动阅读研究领域的创新，同时将框架应用于实际场景中，也可为阅读服务的转型和升级提供新的思路和技术支持。

构建多模态沉浸式智慧阅读模型虽然有很多潜力和优点，但也面临着一些挑战。模型中涉及的自然语言处理、深度学习等技术需要结合实际应用场景进行优化；数据集的规模和质量需要进一步改善；模型的泛化能力和稳定性需要更加深入地研究。未来可以尝试从更多维度对模型的功能进行拓展，以形成一个更为全面的功能框架。随着人工智能技术的不断发展和图书馆服务的不断升级，基于 AI 技术的多模态沉浸式智慧阅读服务将会得到更加广泛的应用和推广。

### 参考文献：

- [1] WALTON G. Smart reading for the smart library: will it transform reading and learning?[J]. New Review of Academic Librarianship, 2015, 21(2):123-125.
- [2] ZHANG Y, CHEN X. Research on the development of wisdom reading under the background of "Internet+"[J]. Library and Information Service, 2015, 59(16):17-22.
- [3] YU S. Research on the construction of wisdom reading service system in public libraries[J]. Library and Information Service, 2017, 61(16):87-93.
- [4] 钟波, 柳青. 人工智能时代下智慧图书馆的构

- 建[J]. 情报理论与实践, 2020, 43(10):136-141.
- [5] 蒋丽萍, 张蓓蓓, 龙晓岚. 智慧图书馆服务模式探析[J]. 图书馆建设, 2020, 40(11):66-69.
- [6] 赵秋莲. 基于智能技术的智慧图书馆建设探究[J]. 现代情报, 2020, 40(6):68-73.
- [7] 郑秀慧, 马靖. 基于大数据的智慧图书馆建设研究[J]. 情报科学, 2019, 37(5):51-56.

梁建春 广西师范大学图书馆讲师。广西桂林, 541004。

(收稿日期:2023-04-07 编校:曹晓文,左静远)