

人工智能在心理健康服务中的应用进展

黄洋洋 周燕玲* 游雨欢

江西中医药大学计算机学院, 江西南昌, 330004;

摘要: 心理健康需求的不断增加与人工智能技术的快速进步, AI驱动的心理支持系统成为推动心理服务普及化、智能化的重要方式。本文梳理人工智能用于心理健康领域的主要技术路径, 重点阐释自然语言处理、多模态情绪识别、基于机器学习的个性化干预机制等核心技术的实现逻辑, 结合实际案例分析其应用场景和使用效果。现有系统仍存在伦理标准、跨领域协作和文化适配等现实问题, 文章提出实现人工智能在心理服务中的深度整合, 需要在技术创新和伦理规则之间寻找平衡, 构建既有专业可信度又有文化适应性的智能心理服务体系, 为心理健康服务的持续发展提供理论支持。

关键词: 人工智能; 心理支持系统; 心理健康

DOI: 10.69979/3041-0673.25.08.006

引言

社会经济结构快速演变和生活节奏不断加快, 公众心理健康问题逐渐成为社会焦点。传统心理服务受限于专业人员不足、服务覆盖范围有限、响应滞后等问题, 难以应对多样化需求。人工智能凭借高效率与低成本可实现全天候响应和低门槛服务接入, 正在重塑心理服务模式。本文综合认知行为疗法和情感计算理论, 梳理人工智能赋能心理健康服务的关键技术, 探讨其在实际场景中的作用与局限, 分析当前面临的模型可控性、伦理治理、跨专业协作和文化适配等核心问题, 为构建专业可靠、技术可控、本土适配性强的智能心理服务体系提供决策参考。

1 人工智能心理支持的相关理论

1.1 认知行为疗法

认知行为疗法认为情绪和行为问题多源于认知偏差, 若能识别并纠正这些偏差, 就帮助个体改善心理状态^[1]。

1.2 情感计算

情感计算使计算机能够识别、理解、处理和模拟人类情绪, 融合了计算机科学、心理学、认知科学等多学科知识^[2]。

1.3 解释水平理论

解释水平理论研究心理距离对信息加工方式的影响^[3]。人工智能心理支持使用 CLT 理论在人机交互过程

中动态调整社会距离, 模拟人际关系发展过程, 缓解用户对“机器支持”的心理隔阂^[4]。

2 人工智能心理支持系统的核心技术

2.1 自然语言处理技术

自然语言处理技术赋予计算机理解和生成自然语言的能力, 是实现人机语言交互的关键基础^[5]。

2.2 多模态情绪识别技术

多模态情绪识别综合利用语音、面部表情与生理信号等多种来源的信息, 提升人工智能系统对情绪识别的精度。

2.3 机器学习技术

机器学习是实现计算机从数据中自动学习规律、建立模型并做出预测或决策的技术框架, 学习方式有有监督、无监督以及深度学习等多种类型^[6]。

3 人工智能心理干预系统的技术实践

3.1 基于规则与检索干预系统

规则与检索型系统以结构化语言模板为核心, 按照预设路径为用户提供干预建议, 代表系统如 Woebot 和 Tess。

Woebot 使用结构化语言引导用户识别自身的非理性想法, 逐步完成认知重构过程^[2]。相比生成型对话系统, Woebot 依托规则模板驱动, 虽在语言灵活性上略显局限, 但有效避免生成不当内容的风险。

Tess 系统采用关键词匹配和脚本检索技术, 为用户

提供即使的文字支持和认知策略建议。短期干预试验显示, Tess 用户在自评问卷“情绪管理效能”平均提升 17%^[7]。不过也有部分用户表示系统回应风格较为单一, 缺乏个性化交流的空间, 在情绪波动较大的状态下干预效果有一定局限。

3.2 基于生成模型干预系统

生成型系统强利用深度神经网络动态生成干预内容, 强调语言表达自然度与个性化, 适用于更复杂的情绪交互场景。

AMN 模型引入情绪记忆机制, 提升情绪连贯性。在心理咨询深度神经网络任务中, AMN 系统在“语言贴切性”与“情绪理解度”两项指标较传统 Transformer 模型提升了 19%与 22%^[8]。

Sharma 等人提出基于大语言模型构建的认知重构辅助系统, 通过自动生成多种积极认知表达, 帮助用户调整负面思维^[9]。用户普遍认为系统“语言自然, 富有启发性”, 也指出偶尔出现“乐观失真”等问题。

3.3 基于人机协同干预系统

人机协同系统结合了心理学干预策略与个性化交互设计, 尤其注重用户信任与文化适配, 增强用户信任与行为转化意愿。

Sharma 等人提出将普通支持性语句重写为具有 MI 风格的表达方式, 采用引导性提问和鼓励性语言促发用户表达下一步意图。有关实验数据显示用户表达行为计划的比例从 34%提升至 66%, 在具有初步表达意愿的用户群体中效果显著^[10]。

在跨文化干预需求日益增长的背景下, Mahmood 团队开发的低资源语言同理表达系统展示出文化适应潜力。系统融合语义控制与多任务同理生成机制, 用户对系统在“语言亲和力”与“情绪理解”两维度的打分近乎满分, 明显优于传统生成模型, 为多语种心理 AI 系统建设提供实践范式^[11]。

4 人工智能心理支持系统的挑战

4.1 伦理治理

人工智能心理支持系统在运行中涉及大量个人敏感数据, 一旦缺乏规范的数据管理机制极易引发敏感信息泄露风险。

建立心理干预分级制度对于规避这些风险具有重要意义。轻度情绪波动时, 系统应具备自主响应能力,

提供标准化调适方案满足用户即时的心理需求, 减少人工资源占用。当检测到中等程度心理健康问题, 系统需融入人机协同工作机制, 由专业心理咨询师跟进评估干预效果。对于严重心理危机情况, 系统必须配备风险预警机制, 第一时间发出警报并自动启动紧急转介流程。

设立心理健康 AI 伦理审查委员来巩固伦理治理。委员会在系统上线前审查系统的设计规范、数据处理流程及干预方式, 确保系统建设符合伦理与法律标准。

4.2 跨学科协同

现阶段大多数 AI 心理系统仍由计算机工程团队主导研发, 心理学专业力量往往在后期才参与到系统设计, 导致系统在情绪识别、语言表达与干预机制设计方面与真实心理机制存在偏差。有效的系统研发应建立计算机科学与心理学之间的协同机制, 例如“人工审查+AI 生成”混合干预架构、联合人才培养计划及早期共建评估流程, 使干预逻辑兼具技术合理性, 又符合心理干预的专业规范。打通技术与心理两端的信息壁垒, 是提升系统科学性与用户体验的关键。

4.3 文化适配

文化背景影响着系统的使用效果和用户接受度。现阶段广泛应用的心理健康 AI 系统大多以西方心理学理论为基础, 难以完全适应中国用户的情绪表达特点和心理模式, 降低了系统干预效果和信任度。

在提高系统文化适配性的实践中, 将传统中医“情志理论”融入 AI 干预模型的尝试取得积极成果。中医情志理论认为情绪具有意向性, 即情绪不仅体现为主观感受, 还通过眼动、生理反应和面部表情等外显指标表现出来。悲伤时注视次数增多、皮肤电导变化周期长, 而喜悦时注视次数减少、反应更快^[12]。中西方用户在信任构建上也存在文化差异, 西方用户更关注系统的透明度与个性化定制, 信任构建更依赖于系统自身表现。而中国用户更倾向于参考社会评价、权威机构认证。

5 结论与展望

人工智能心理支持系统正逐步变为推动心理健康服务普及、智能化和精准化的一股重要力量。本文以心理学理论为依据, 梳理了当前应用中的核心技术路径, 结合实际案例分析其在服务优化中的实践价值, 并指出该领域在模型可控性、伦理规范、跨学科协作与文化适切中面临的问题。未来应加强心理理论在系统设计中的融

合,推动干预策略标准化与表达方式本土化,构建既安全可信又具文化亲和力的智能心理服务体系,为复杂多样的心理健康需求提供长期支持。

参考文献

- [1] BECK, A. T. Thinking and depression. II. Theory and therapy[J]. Archives of General Psychiatry, 1964, 10: 561 - 571.
- [2] 罗森林,潘丽敏. 情感计算理论与技术[J]. 系统工程与电子技术,2003,25(7):905-909.
- [3] 黄俊,李晔,张宏伟. 解释水平理论的应用及发展[J]. 心理科学进展,2015,23(01):110-119.
- [4] 林雅萍,郑浩然,何忆君,冯晶. AI交互质量与用户接受度:心理距离和信任的链式中介作用[J]. 心理科学进展,2025,15(2):47 - 57. <https://doi.org/10.12677/ap.2025.152060>.
- [5] Hirschberg J, Manning C D. Advances in natural language processing[J]. Science, 2015, 349: 261-266.
- [6] Calvo R A, Milne D N, Hussain M S, et al. Natural language processing in mental health applications using non-clinical texts[J]. Natural Language Engineering,2017,23(5):649-685.
- [7] Fulmer R, Joerin A, Gentile B, et al. Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: randomized controlled trial[J]. JMIR mental health, 2018, 5(4): e9782.
- [8] Zhang Z, Zhang T, Li H. Affective Dialogue Generation with Affective Memory Network[C]//

Proceedings of the Findings of the Association for Computational Linguistics: ACL 2023. Toronto, Canada: ACL, 2023: 4335 - 4351. DOI:10.18653/v1/2023.findings-acl.334.

[9] Sharma A, Ghosh S, Dey A. Assisting Cognitive Restructuring with Large Language Models: Towards AI-Augmented Mental Health Support[C]// Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. Honolulu, USA: ACM, 2024. DOI:10.1145/3613904.3642761.

[10] Sharma A, Pathak R, Danescu-Niculescu-Mizil C. Rewriting Supportive Responses to Online Mental Health Posts to Align with Motivational Interviewing[C]// Proceedings of the BioNLP Workshop 2022. Dublin, Ireland: ACL, 2022: 268 - 278. DOI:10.18653/v1/2022.bionlp-1.27.

[11] Mahmood A, Jahanbani M, Riazi A. Empathetic Response Generation for Mental Health Support in Low-Resource Settings[C]// Proceedings of AbjadNLP 2025. Tehran, Iran: ACL, 2025: 76 - 86. DOI:10.18653/v1/2025.abjadnlp-1.8.

[12] 陈玉霏. 基于中医情志理论的悲喜情绪识别模式研究[D]. 广东:广州中医药大学,2017.

作者简介:黄洋洋(2002-12-),女,汉族,浙江温州人,江西中医药大学2022级医学信息工程班学生。
*通信作者简介:周燕玲(1977-),女,汉族,河南许昌人,大学本科,副教授,主要从事计算机应用技术研究。