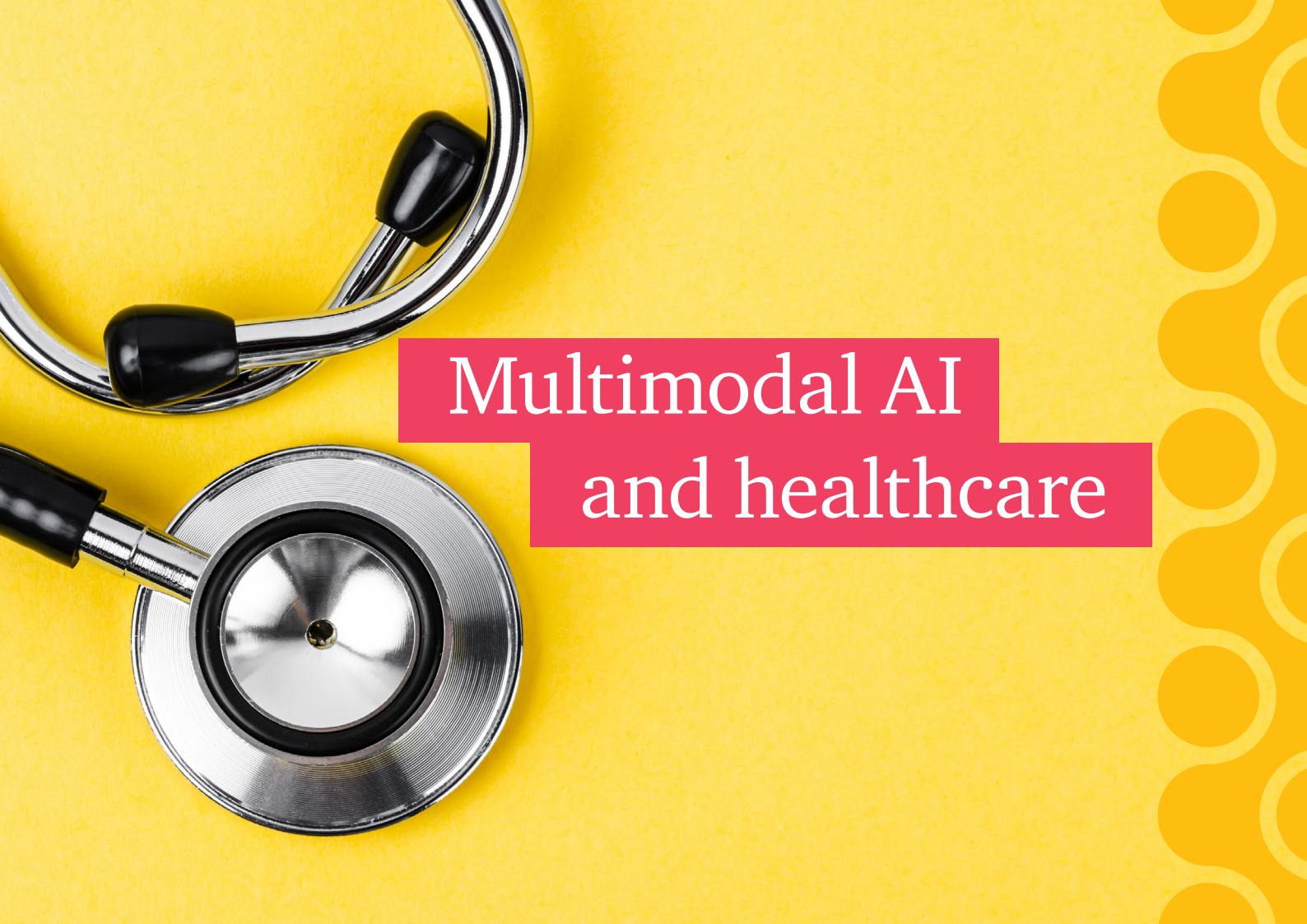




Multimodal AI and healthcare

Overview

As humans, we feel the world around us in a multimodal manner – we can feel texture, hear sounds, see objects, smell various scents and taste food. Standard AI systems are normally unimodal, which implies that they are trained to do one job, like images or language processing. Such systems can detect words or images using a single dataset. Though it is simple to handle a single source of data from the perspective of a standard process, its software lacks context and supportive information to minimise the unstructured data. The ability to analyse multimodal inputs like humans is a prerequisite for advanced AI systems.

Multimodal AI refers to AI systems that can analyse and recognise data from numerous modalities or sources. The multimodal AI model offers a wide range of applications in industries such as healthcare, finance and entertainment.

Multimodal AI can be leveraged to convey these sensory experiences more effectively, thus increasing the accuracy of communication. Moreover, it can facilitate simultaneous sight and hearing, which is similar to the way humans communicate with each other.

Natural language processing (NLP) is one example of a multimodal paradigm, which integrates text and speech recognition to enable more accurate and natural language exchanges between humans and machines. Another example is picture recognition, which can be enhanced by combining information from several modalities like text and voice.

Benefits of multimodal AI in healthcare

- a. **Faster and more accurate diagnostics:** Multimodal AI can achieve more accuracy and robustness than unimodal AI by combining input from several modalities.
- b. **Reliability in the face of noise:** Multimodal AI can be more resistant to noise and fluctuation in incoming data by incorporating information from different modalities.
- c. **Improved user experience:** By providing numerous ways for consumers to interact with the system, multimodal AI can improve the user experience.
- d. **Effective resource management:** By allowing the system to focus on the most relevant information from each modality, multimodal AI can help to make better use of computing and data resources.
- e. **Enhanced interpreting:** By offering several sources of knowledge that can be utilised to explain the system's output, multimodal AI can help in increasing interpretability.

Multimodal AI technology stack

NLP: Allows speech recognition, thus enabling the system to understand and transcribe spoken language.

Computer vision technologies: Enable the system to contextualise actions, objects, and people by analysing and decoding complicated visual input, thus facilitating video and picture recognition.

Textual analysis: Enables the system to comprehend written material, including sentiment analysis and language translation.

Technologies for speed processing and data mining: These will facilitate faster calculations in real time.



How does a multimodal system work?

Any therapy should cover all facets of the human condition. The human mind is a complex organ which plays a key role in performing and assessing various tasks/actions. The importance of the mind in both a normal and ill state has been long established. In the upcoming decade, AI systems have tremendous potential to advance healthcare. Healthcare operations will soon be able to rely extensively on AI and ML as core tools to increase effectiveness of treatment. We've elaborated the different phases to run such a system below.

Data acquisition: The multimodal system collects data from different sources like electronic health records (EHRs), medical imaging devices (MRI, CT scans), wearable devices (fitness trackers, smart watches) and patient-reported outcomes (symptom questionnaires).

Data integration: This phase involves incorporating clinical data, imaging results, vital signs and patient-reported information into a single patient profile, and the system merges the gathered data into a unified platform.

Data pre-processing: To ensure data quality, consistency and compatibility, pre-processing procedures are applied to the collected data. As part of this, data formats must be standardised, noise or mistakes have to be eliminated, and values need to be normalised for investigation.

Data analysis: The integrated data is analysed using sophisticated algorithms and AI approaches. For instance, ML algorithms can examine medical images to spot anomalies, NLP algorithms extract crucial details from clinical notes, and predictive modelling can estimate the risk of whether a disease will advance.

Clinical decision support: To aid in clinical decision making, the multimodal system generates observations and suggestions. It provides healthcare professionals with necessary details, including possible diagnoses, options for treatment and customised treatment strategies based on the investigated data.

Communication and collaboration: The system makes it feasible for healthcare professionals to collaborate and communicate effectively. It gives doctors access to a common platform where they can share knowledge and exchange treatment plans and additionally enhance care coordination and interdisciplinary cooperation.

Patient monitoring and follow-up: Through the multimodal system, healthcare providers can remotely monitor patients' health status. Real-time data on vital signs, activity levels and medication adherence are gathered via wearable technology and home-based sensors, enabling proactive interventions and remote consultations as per requirement.

Patient engagement and learning: The multimodal system educates and engages patients by offering interactive tools, individualised health information and instructional resources. Patients can access their health information, monitor their progress, get reminders for appointments or medication and access condition-specific instructions.

Research and community health: Data from the multimodal system can be aggregated and anonymised for use in population health studies and medical research. Large datasets can be analysed by researchers to find illness patterns, assess the efficacy of treatments and support evidence-based practices.



How PwC can help

PwC has a strong structure in place to assist you in comprehending the influence of technology on numerous business domains. Its experienced and dedicated workforce can help clients by building any application according to the required specifications.

Moreover, the firm provides highly effective application moderation services which can work against distribution shifts as well as other runtime issues.

In addition, PwC has an experienced workforce in cloud transformation services which can help with computational resource requirements, thus reducing the cost for training and running the services.

The firm can help you to identify and prioritise multimodal AI use cases based on their potential to create business value.



About PwC

At PwC, our purpose is to build trust in society and solve important problems. We're a network of firms in 152 countries with over 328,000 people who are committed to delivering quality in assurance, advisory and tax services. Find out more and tell us what matters to you by visiting us at www.pwc.com.

PwC refers to the PwC network and/or one or more of its member firms, each of which is a separate legal entity. Please see www.pwc.com/structure for further details.

© 2023 PwC. All rights reserved.

Contact us

Ashootosh Chand

Partner, Digital and Emerging Technologies

PwC India

ashootosh.chand@pwc.com

Indrojeet Bhattacharya

Managing Director, Digital and Emerging Technologies – Web 3.0

PwC India

indrojeet.bhattacharya@pwc.com

Krishnendu Sadhu

Manager, Digital and Emerging Technologies

PwC India

krishnendu.sadhu@pwc.com

Contributors

Krishnendu Sadhu

Editorial support

Rashi Gupta

Design

Shipra Gupta

pwc.in

Data Classification: DC0 (Public)

In this document, PwC refers to PricewaterhouseCoopers Private Limited (a limited liability company in India having Corporate Identity Number or CIN : U74140WB1983PTC036093), which is a member firm of PricewaterhouseCoopers International Limited (PwCIL), each member firm of which is a separate legal entity.

This document does not constitute professional advice. The information in this document has been obtained or derived from sources believed by PricewaterhouseCoopers Private Limited (PwCPL) to be reliable but PwCPL does not represent that this information is accurate or complete. Any opinions or estimates contained in this document represent the judgment of PwCPL at this time and are subject to change without notice. Readers of this publication are advised to seek their own professional advice before taking any course of action or decision, for which they are entirely responsible, based on the contents of this publication. PwCPL neither accepts or assumes any responsibility or liability to any reader of this publication in respect of the information contained within it or for any decisions readers may take or decide not to or fail to take.

© 2023 PricewaterhouseCoopers Private Limited. All rights reserved.

SG/August 2023-M&C 30297