मल्टीमॉडल ए.आई का उद्भव

प्रिलिम्स के लिये:

मल्टीमोडल AI उद्भव, AI(कृत्रिम बुद्धिमत्ता), ह्यूमन-लाइक कॉग्निशन, ओपन AI चैट जीपीटी, गूगल जेमिनी मॉडल।

मेन्स के लिये:

मल्टीमोडल AI का उद्भव और उसके निहितार्थ, मल्टीमॉडल AI का विकास और उसके अनुप्रयोग तथा दैनंदिन के जीवन पर प्रभाव

स्रोत: द हिंदू

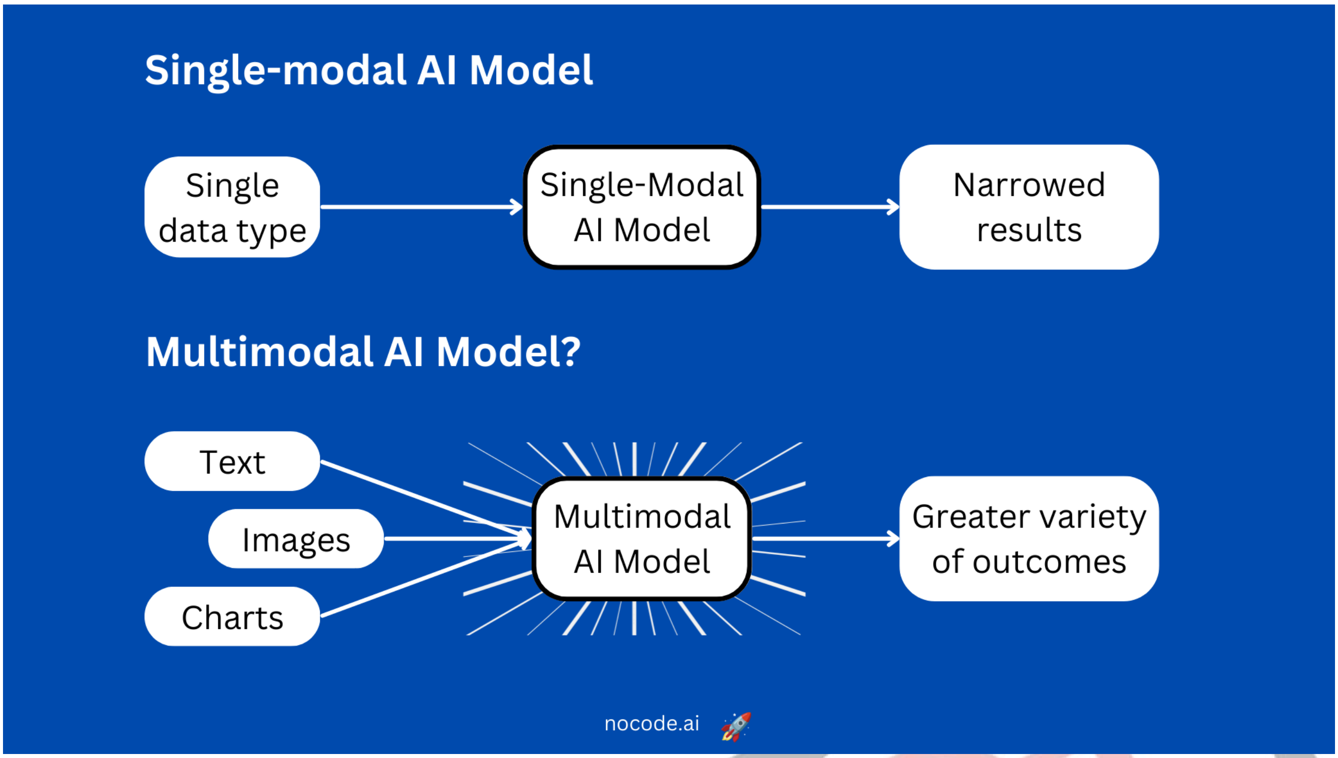
चर्चा में क्यों?

AI (कृत्रिम बुद्धिमत्ता) ने मल्टीमोडल सिस्टम की दिशा में एक आदर्श परिवर्तन किया है, जो लोगों को टेक्स्ट, छवियों, ध्वनियों और वीडियो के माध्यम से AI के साथ बातचीत करने में सक्षम बनाता है।

- इन प्रणालियों का लक्ष्य विभिन्न प्रकार के **संवेदी** प्रसंस्करण (**Sensory Input)** का उपयोग करके मानव जैसे संज्ञान की नकल करना है।

मल्टीमोडल AI सिस्टम

- **परिचय:**
  - मल्टीमोडल AI कृत्रिम बुद्धिमत्ता को संदर्भित करता है जो वास्तविक दुनिया के मुद्दों के संबंध में**अधिक सटीक पूर्वानुमान, व्यावहारिक निष्कर्ष अथवा निर्णय करने के लिये कई डेटा प्रकारों अथवा मोड को एकीकृत कर सकता है**।
  - वीडियो, ऑडियो, भाषण, चित्र, पाठ और विभिन्न प्रकार के पारंपरिक संख्यात्मक डेटा सेट का उपयोग तथा प्रशिक्षण मल्टीमोडल AI सिस्टम द्वारा किया जाता है।
  - **उदाहरणत:** व्हिस्पर, ओपन-AI का ओपन-सोर्स स्पीच-टू-टेक्स्ट ट्रांसलेशन मोडल, जीपीटी की **वॉयस प्रोसेसिंग क्षमताओं का आधार** है। मल्टीमोडल ऑडियो सिस्टम **समान सिद्धांतों** पर कार्य करते हैं।



- **मल्टीमोडल AI में हालिया विकास:**
    - OpenAI ने अपने GPT-3.5 और GPT-4 मोडल में संवर्द्धन की घोषणा की, जिससे उन्हें छवियों का विश्लेषण करने तथा स्पीच सिंथेसिस में संलग्न होने में सहायता मिली, जिससे उपयोगकर्त्ताओं के साथ अधिक गहन इंटरेक्शन संभव हो सका।
        - यह "गोबी" नाम के एक प्रोजेक्ट पर काम कर रहा है, जिसका लक्ष्य GPT मॉडल से अलग एक नए सिरे से मल्टीमोडल AI सिस्टम बनाना है।
    - **गूगल का जेमिनी मोडल:**
        - इस क्षेत्र में एक अन्य प्रमुख दिग्गज **Google का** नया **मल्टीमोडल लार्ज लैंग्वेज मोडल जो अब तक रिलीज़ नहीं हुआ** है, **Gemini** है।
            - अपने सर्च इंजन और यूट्यूब से छवियों एवं वीडियो के विशाल संग्रह के कारण,**Google को मल्टीमोडल डोमेन में अपने प्रतिद्वंद्वियों पर स्पष्ट बढ़त हासिल** थी।
            - यह अन्य AI प्रणालियों पर अपनी मल्टीमोडल क्षमताओं को तेज़ी से आगे बढ़ाने के लिये अत्यधिक दबाव डालता है।

यूनिमोडल AI की तुलना में मल्टीमॉडल AI के फायदे:

- मल्टीमोडल AI, यूनिमोडल AI के विपरीत **टेक्स्ट, चित्र और ऑडियो जैसे विविध डेटा प्रकारों** का लाभ उठाता है, जो जानकारी का एक समृद्ध प्रतिनिधित्व प्रदान करते हैं।
- यह दृष्टिकोण प्रासंगिक समझ को बढ़ाता है, जिसके परिणामस्वरूप अधिक सटीक अनुमान और सूचित निर्णय सुनिश्चित होते हैं।
- कई तौर-तरीकों से डेटा को फ्यूज़ करके, **मल्टीमोडल AI बेहतर प्रदर्शन, सुदृढ़ता और अस्पष्टता** को प्रभावी ढंग से कार्यान्वित करने की **क्षमता अर्जित करता** है।
- यह **विभिन्न डोमेन में प्रयोज्यता को व्यापक** बनाता है और **क्रॉस-मोडल लर्निंग को सक्षम** बनाता है।
- **मल्टीमोडल ए.आई. डेटा की अधिक समग्र और मानव-जैसी समझ प्रदान करता है** , यह इसके नवीन अनुप्रयोगों तथा जटिल वास्तविक वैश्विक परिदृश्यों की गहन समझ का मार्ग प्रशस्त करता है।

मल्टीमोडल ए.आई.के अनुप्रयोग:

- मल्टीमोडल ए.आई का उपयोग **स्वायत्त ड्राइविंग**, **रोबोटिक्स** और **चिकित्सा सहित विभिन्न क्षेत्रों** में संभव है।
    - उदाहरण के लिये, चिकित्सा क्षेत्र में **सी.टी. स्कैन** द्वारा जटिल डेटासेट का विश्लेषण और आनुवंशिक विविधताओं की पहचान का कार्य, चिकित्सा पेशेवरों के लिये परिणामों की साझाकरण प्रक्रिया को सरल बनाना आदि कार्य महत्त्वपूर्ण हैं।
- गूगल ट्रांसलेट और Meta के Seamless M4T जैसे स्पीच ट्रांसलेशन मॉडल को मल्टीमोडलिटी से लाभ मिलता है, ये सभी मॉडल विभिन्न भाषाओं में अनुवाद सेवाएँ प्रदान करते हैं।
- हाल के इस क्षेत्र में हुए विकासों में मेटा का इमेजबाइंड (ImageBind) प्रमुख है, यह एक मल्टीमोडल प्रणाली है जो **टेक्स्ट**, विज़ुअल डेटा, ऑडियो, तापमान और मूवमेंट रीडिंग को संसाधित करने में सक्षम है।
    - इसमें स्पर्श, गंध, भाषण और MRI मस्तिष्क संकेतों जैसे अतिरिक्त संवेदी डेटा को एकीकृत करने की संभावनाओं पर विचार किया जा रहा है, ताकि भविष्य में ए.आई. प्रणाली को जटिल वातावरण का अनुकरण करने में सक्षम बनाया जा सके।

मल्टीमोडल ए.आई. की चुनौतियाँ:

- **डेटा की मात्रा और भंडारण:**
    - मल्टीमोडल ए.आई. के लिये विविध और विशाल डेटा की आवश्यक होती है जो**डेटा गुणवत्ता, भंडारण लागत एवं अतिरिक्त प्रबंधन के मुद्दों के कारण महंगा और संसाधन-गहन(जिनके लिये व्यापक संसाधनों और ऊर्जा की आवश्यकता होती है) है।**
- **संदर्भ और बारीकियों की समझ:**
    - **एक समान इनपुट के विभिन्न सूक्ष्म अर्थों की समझ तैयार करने के लिये AI को प्रशिक्षित करने का कार्य** विशेष रूप से भाषाओं अथवा संदर्भ आधारित अर्थों वाली अभिव्यक्तियों में स्वर, चेहरे के भाव जैसे अन्य प्रासंगिक संकेतों के बिना चुनौतीपूर्ण साबित होता है।
- **सीमित और अपूर्ण डेटा:**
    - असीमित और **आसानी से पहुँच योग्य डेटा समूह की उपलब्धता एक चुनौती** है। सार्वजनिक डेटा समूह सीमित, महँगे या एकत्रीकरण समस्याओं से ग्रस्त हो सकते हैं, जिससे AI मोडल प्रशिक्षण में डेटा अखंडता और पूर्वाग्रह प्रभावित हो सकते हैं।
- **गुम डेटा प्रबंधन:**
    - एकाधिक स्रोतों से **डेटा पर निर्भरता के परिणामस्वरूप AI में खराबी हो सकती** है या किसी भी डेटा स्रोत की गलत व्याख्या हो सकती है, जिससे AI प्रतिक्रिया में अनिश्चितता उत्पन्न हो सकती है।
- **निर्णय लेने की जटिलता:**
    - मल्टीमोडल AI में तंत्रिका नेटवर्क की व्याख्या करना जटिल और चुनौतीपूर्ण हो सकता है, **जिससे यह समझना मुश्किल हो जाता है** कि AI डेटा का मूल्यांकन किस प्रकार करता है तथा निर्णय कैसे लेता है। पारदर्शिता की यह कमी डिबगिंग और पूर्वाग्रह उन्मूलन प्रयासों में बाधा बन सकती है।

निष्कर्ष:

- मल्टीमोडल AI सिस्टम का आगमन कृत्रिम बुद्धिमत्ता के क्षेत्र में एक महत्त्वपूर्ण प्रगति का प्रतिनिधित्व करता है।
- इन प्रणालियों में विभिन्न उद्योगों में क्रांति लाने, मानव-कंप्यूटर इंटरैक्शन को बढ़ाने और जटिल वास्तविक दुनिया की समस्याओं का समाधान करने की क्षमता है।
- जैसे-जैसे AI का विकास जारी है, मल्टीमोडैलिटी कृत्रिम सामान्य बुद्धिमत्ता प्राप्त करने और AI अनुप्रयोगों की सीमाओं का विस्तार करने में महत्त्वपूर्ण भूमिका निभाने के लिये तैयार है।

PDF Refernece URL: https://www.drishtiias.com/hindi/printpdf/emergence-of-multimodal-ais