ChatGPT और संबद्ध नैतिक चिंताएँ

प्रिलिम्स के लिये:

कृत्रिम बुद्धिमत्ता, ChatGPT, ह्यूमन फीडबैक रीइन्फोर्समेंट लर्निंग (RLFH), फिशिंग ईमेल, प्लैजरिज़्म।

मेन्स के लिये:

ChatGPT का महत्त्व और इससे जुड़ी नैतिक चिंताएँ।

चर्चा में क्यों?

हाल ही में **OpenAI (AI अनुसंधान और परिनियोजन कंपनी)** ने अपनी क्षमता का परीक्षण करने के लिये उपयोगकर्त्ताओं हेतु अपना सबसे हालिया और शक्तिशाली **AI चैटबॉट**, **ChatGPT** शुरू किया है।

ChatGPT:

- **परिचय:**
  - ChatGPT **जनरेटिव प्री-ट्रेन ट्रांसफार्मर (GPT)** का एक प्रकार है जो **OpenAI द्वारा विकसित बड़े पैमाने पर तंत्रिका नेटवर्क-आधारित भाषा मॉडल है।**
  - GPT मॉडल को मानव-समान पाठ करने हेतु बड़ी मात्रा में पाठ डेटा पर प्रशिक्षित किया जाता है।
  - यह विषयों की एक विस्तृत शृंखला के लिये प्रतिक्रियाएँ उत्पन्न कर सकता है, जैसे प्रश्नों का उत्तर देना, स्पष्टीकरण प्रदान करना और संवाद में शामिल होना।
  - ChatGPT "अनुवर्ती प्रश्नों" का उत्तर दे सकता है और **"अपनी गलतियों को स्वीकार कर सकता है, गलत धारणाओं को चुनौती दे सकता है, साथ ही अनुचित अनुरोधों को अस्वीकार कर सकता है।"**
  - चैटबॉट को **रीइन्फोर्समेंट लर्निंग फ्रॉम ह्यूमन फीडबैक (RLHF)** का उपयोग करके भी प्रशिक्षित किया गया था।
- **सीमाएँ:**
  - चैटबॉट में स्पष्ट रूप से **नस्लीय और लैंगिक पूर्वाग्रह देखने को मिले** जो लगभग सभी AI मॉडल की समस्या है।
  - चैटबॉट के उत्तर व्याकरणिक रूप से सही होते हैं, यह अच्छी तरह से पढ़ भी सकता है, हालाँकि कुछ शोधकर्त्ताओं के अनुसार **इसमें संदर्भ एवं सार को समझने में समस्या है और यह काफी हद तक सच है।**
  - ChatGPT कभी-कभी गलत जानकारी प्रदान करता है **और वर्ष 2021 से पहले हुई वैश्विक घटनाएँ इसके ज्ञान की सीमा है।**

ChatGPT से संबद्ध नैतिक चिंताएँ:

- **खतरनाक कोडिंग:** कुछ उपयोगकर्त्ता चैटबॉट की क्षमता का गलत उपयोग कर रहे हैं और इसके बारे में दावे भी किये जाते रहते हैं।
- **फिशिंग ईमेल की उत्पत्ति: फिशिंग ईमेल या दुर्भावनापूर्ण कोड लिखने के अनुरोधों को अस्वीकार करने के लिये** ChatGPT की स्थापना की गई, **लेकिन वास्तव में यह उत्कृष्ट फिशिंग ईमेल की उत्पत्ति में मदद कर रहा है।**
- **पक्षपाती डेटा/सूचना का निर्माण:** उत्पन्न किये गए कोड में **पूर्वाग्रह की संभावना भी चिंता का विषय है।**
- **रोज़गार के अवसर को कम करना:** एक चिंता यह है कि **कोड जनरेटर के उपयोग** से मानव प्रोग्रामर के **रोज़गार का नुकसान** हो सकता है।
- **प्लेगरिज़्म चोकपॉइंट:** शिक्षाविदों में **साहित्यिक चोरी (प्लेगरिज़्म)** कोई नया मुद्दा नहीं है, लेकिन ChatGPT ने बता दिया है कि मूल लेखन का उत्पादन करने के लिये AI का उपयोग कैसे किया जाता है। परिणामस्वरूप चोरी की गई जानकारी की पहचान करना मुश्किल है। शिक्षकों एवं शिक्षाविदों ने भी लिखित असाइनमेंट पर ChatGPT के प्रभाव पर चिंता व्यक्त की है।

आगे की राह

- **संदर्भ को समझने और प्रतिक्रिया देने की अपनी क्षमता में सुधार करना,** जैसे किसी विशिष्ट विषय पर संवाद को बनाए रखने में सक्षम होना या

किसी संदेश के स्वर तथा मंशा को समझना।

- अधिक सटीक और सूचनात्मक प्रतिक्रियाएँ उत्पन्न करने के लिये **अपने ज्ञान के आधार एवं क्षमता को बढ़ाना।**
- **अधिक उन्नत भाषा निर्माण क्षमताओं का विकास करना** , जैसे कि सुसंगत एवं प्रेरक निबंध लिखने में सक्षम होना या रचनात्मक एवं आकर्षक कल्पना करना।
- विभिन्न भाषाओं और बोलियों को **समझने तथा प्रतिक्रिया देने की अपनी क्षमता में सुधार करना।**
- **अधिक उन्नत संवादात्मक क्षमताओं का विकास करना** जैसे कि कई कार्यों को संभालना और अधिक जटिल वार्तालापों में संलग्न होना।

यूपीएससी सिविल सेवा परीक्षा, विगत वर्ष के प्रश्न

प्रश्न. विकास की वर्तमान स्थिति के साथ आर्टिफिशियल इंटेलिजेंस प्रभावी रूप से निम्नलिखित में से क्या कर सकता है? (2020)

औद्योगिक इकाइयों में बिजली की खपत कम करना

सार्थक लघु कथाएँ और गीत रचना

रोग निदान

पाठ से वाक् रूपांतरण

विद्युत ऊर्जा का वायरलेस ट्रांसमिशन

नीचे दिये गए कूट का प्रयोग कर सही उत्तर चुनिये:

(a) केवल 1, 2, 3 और 5

(b) केवल 1, 3 और 4

(c) केवल 2, 4 और 5

(d) 1, 2, 3, 4 और 5

उत्तर: (b)

स्रोत: द हिंदू

PDF Refernece URL: https://www.drishtiias.com/hindi/printpdf/chatgpt-and-associated-ethical-concerns