सर्वम टोकन फैक्ट्री: AI आत्मनिर्भरता के अलग आयाम

समाचार-संदर्भ: 16 नवंबर 2026 | प्रकाशन: 19 नवंबर 2026
मुख्य परीक्षा: GS-III, विज्ञान-प्रौद्योगिकी और डिजिटल अर्थव्यवस्था



मौलिक प्रतीकात्मक चित्र: pcsmantra.com

▼ इस लेख में

- खबरों में क्यों
- इंफरेंस क्या है
- मॉडल और उसे चलाने की जगह
- आत्मनिर्भरता को कैसे परखें
- अवसर और सावधानी
- आगे की राह
- प्रिलिम्स परीक्षा अभ्यास प्रश्न
- मुख्य परीक्षा अभ्यास प्रश्न
- अक्सर पूछे जाने वाले प्रश्न (FAQ)
- स्रोत

खबरों में क्यों

भारतीय AI कंपनी सर्वम ने 16 नवंबर 2026 को Token Factory पेश किया। कंपनी के अनुसार, यह भारत में स्थित GPU समूहों से उसके अपने तथा अन्य संस्थानों के ओपन-वेट मॉडल उपलब्ध कराने वाला इंफरेंस प्लेटफॉर्म है। इसका महत्व किसी एक चैटबॉट से आगे है: भारत में मॉडल बनाना, विदेशी मॉडल को स्थानीय कंप्यूटरों पर चलाना और उपयोगकर्ता के डेटा पर नियंत्रण रखना, तीन अलग प्रश्न हैं। [1]

इंफरेंस क्या है

प्रशिक्षण में बहुत से उदाहरणों से मॉडल के पैरामीटर सीखे जाते हैं। **इंफरेंस** में प्रशिक्षित मॉडल किसी नए अनुरोध का उत्तर तैयार करता है। **GPU** ऐसे गणना-चिप हैं जो AI के समानांतर गणितीय काम में उपयोगी हैं। कंपनी ने Token Factory को इंफरेंस स्टैक कहा है, नया अकेला मॉडल नहीं।

API दो सॉफ्टवेयरों के बीच अनुरोध और उत्तर भेजने का तय तरीका है। सर्वम का दावा है कि उसके एक OpenAI-compatible API से कई मॉडल इस्तेमाल किए जा सकते हैं। Compatible का अर्थ यहां इंटरफेस के अनुरूप होना है, किसी दूसरी कंपनी का स्वामित्व या समर्थन नहीं। [1]

यह भी पढ़ें: **भारत का पहला आर्बिटल डेटा सेंटर उपग्रह: Pixxel-Sarvam की 'पाथफाइंडर' योजना**

मॉडल और उसे चलाने की जगह

लॉन्च विवरण में Sarvam 105B के साथ GLM, DeepSeek, Kimi और Gemma परिवारों के मॉडल सूचीबद्ध हैं। कंपनी इन्हें भारत के GPU समूहों से उपलब्ध कराने की बात कहती है। यह विवरण मॉडल की उत्पत्ति और उसे चलाने की जगह में फर्क दिखाता है। किसी विदेशी मूल के मॉडल को भारत में होस्ट करने से वह भारतीय-प्रशिक्षित मॉडल नहीं बन जाता। [1]

ओपन-वेट का अर्थ मॉडल के सीखे हुए पैरामीटर उपलब्ध होना है। इससे हर मॉडल का पूरा प्रशिक्षण डेटा खुला होने या बिना शर्त उपयोग की अनुमति होने का निष्कर्ष नहीं निकलता। उपयोग की शर्तें संबंधित लाइसेंस से तय होती हैं।

आत्मनिर्भरता को कैसे परखें

मॉडल क्षमता: क्या घरेलू संस्थान भारतीय भाषाओं और जरूरतों के अनुसार मॉडल विकसित और सुधार सकते हैं?

गणना क्षमता: क्या आवश्यक चिप, बिजली, नेटवर्क और संचालन-कौशल भरोसेमंद ढंग से उपलब्ध हैं?

डेटा शासन: अनुरोध, लॉग और आउटपुट कहाँ रखे जाते हैं, किसे उन तक पहुँच है और कितने समय तक उन्हें रखा जाता है? केवल स्थानीय होस्टिंग का दावा इन सभी सवालों का उत्तर नहीं है।

चुनाव की स्वतंत्रता: कई मॉडलों में बदलने की सुविधा एक प्रदाता पर निर्भरता कम कर सकती है। फिर भी API, लाइसेंस और सेवा-शर्तों की जाँच जरूरी है। ये विश्लेषण के मानदंड हैं, इस प्लेटफॉर्म के लिए प्रमाणित सुरक्षा-निष्कर्ष नहीं।

अवसर और सावधानी

स्थानीय इंफरेंस सेवा से किसी छोटे उद्यम को अपना बड़ा कंप्यूटिंग ढाँचा लगाए बिना AI अनुप्रयोग बनाने का रास्ता मिल सकता है। भारतीय भाषाओं और सार्वजनिक सेवाओं के लिए घरेलू क्षमता का विकास भी नीति का घोषित उद्देश्य है। PIB का फरवरी 2026 विवरण सर्वम को भारत के स्वदेशी, बहुभाषी AI परिवेश से जोड़ता है। इसे नवंबर के किसी विशिष्ट व्यावसायिक प्लेटफॉर्म को सरकारी सुरक्षा-प्रमाणन या सरकारी वित्त मिलने का प्रमाण नहीं मानना चाहिए। [2]

कंपनी के लागत और गति संबंधी आंकड़े उसके अपने परीक्षण हैं। अलग मॉडल, अनुरोध की लंबाई, उपयोगकर्ता संख्या और परीक्षण-पद्धति के परिणाम सीधे नहीं मिलाए जा सकते। इस लेख में उन्हें स्वतंत्र प्रदर्शन-रैंकिंग नहीं बनाया गया है। [1]

आगे की राह

AI नीति में घरेलू अनुसंधान और खुले वैश्विक मॉडल के उपयोग को अलग-अलग परखा जा सकता है। संवेदनशील काम के लिए डेटा-प्रवाह, स्वतंत्र सुरक्षा-परीक्षण, लाइसेंस, खरीद-शर्तों और मॉडल बदलने की क्षमता की जाँच जरूरी है। मूल्यांकन का प्रश्न केवल यह नहीं होना चाहिए कि मॉडल किस देश में बना, बल्कि यह भी कि प्रणाली पर वास्तविक नियंत्रण किसका है और उसकी जवाबदेही कैसे तय होती है।

PCSMANTRA.COM

प्रिलिम्स परीक्षा अभ्यास प्रश्न

प्रश्न 1

AI प्लेटफॉर्म के संदर्भ में निम्नलिखित कथनों पर विचार कीजिए:

कथन 1: इंफरेंस में प्रशिक्षित मॉडल नए इनपुट पर आउटपुट तैयार करता है।

कथन 2: भारत में होस्ट किया गया हर मॉडल अनिवार्यतः भारत में प्रशिक्षित होता है।

कथन 3: API दो सॉफ्टवेयरों के बीच अनुरोध-उत्तर का तय तरीका हो सकता है।

उपर्युक्त कथनों में से कौन-से सही हैं?

(a) केवल 1 और 3

(b) केवल 1 और 2

(c) केवल 2 और 3

(d) 1, 2 और 3

▼ उत्तर और व्याख्या

सही उत्तर: (a)

कथन 1 सही है: प्रशिक्षण और इंफरेंस अलग चरण हैं।

कथन 2 गलत है: होस्टिंग का स्थान मॉडल के प्रशिक्षण या उत्पत्ति का प्रमाण नहीं है।

कथन 3 सही है: API सॉफ्टवेयरों के परस्पर संवाद का इंटरफेस है।

PCSMANTRA.COM

प्रश्न 2

Token Factory के लॉन्च विवरण का सबसे उपयुक्त अर्थ क्या है?

(a) केवल एक नया सोशल मीडिया मंच

(b) कई AI मॉडलों की इंफरेंस सेवा के लिए एक प्लेटफॉर्म

(c) सभी AI मॉडलों के प्रशिक्षण डेटा का सरकारी भंडार

(d) केवल चिप निर्माण की फैक्ट्री

▼ उत्तर और व्याख्या

सही उत्तर: (b)

कंपनी इसे अपना तथा अन्य ओपन-वेट मॉडल उपलब्ध कराने वाला इंफरेंस स्टैक बताती है। नाम में Factory होने से यह चिप-निर्माण संयंत्र नहीं बन जाता।

प्रश्न 3

ओपन-वेट मॉडल के बारे में कौन-सा निष्कर्ष सबसे उचित है?

(a) उस पर कोई लाइसेंस लागू नहीं होता

(b) उसका हर प्रशिक्षण उदाहरण सार्वजनिक होता है

(c) उसके सीखे हुए पैरामीटर उपलब्ध हो सकते हैं; उपयोग की शर्तें फिर भी देखनी होती हैं

(d) वह अनिवार्यतः सरकारी स्वामित्व में होता है

▼ उत्तर और व्याख्या

सही उत्तर: (c)

वेट उपलब्ध होने और पूर्ण डेटा अथवा बिना शर्त उपयोग-अधिकार मिलने में अंतर है। लाइसेंस और सेवा-शर्तें अलग से जाँचनी चाहिए।

मुख्य परीक्षा अभ्यास प्रश्न

“AI आत्मनिर्भरता को केवल मॉडल के मूल देश से नहीं मापा जा सकता।” स्थानीय इंफरेंस प्लेटफॉर्म के संदर्भ में परीक्षण कीजिए। (150 शब्द, 10 अंक)

PCSMANTRA.COM

उत्तर की रूपरेखा

भूमिका में:

AI आत्मनिर्भरता का अर्थ अनुसंधान, गणना और डेटा-निर्णयों पर पर्याप्त क्षमता और नियंत्रण है। मॉडल की उत्पत्ति इसका एक आयाम है, पूरा पैमाना नहीं।

मुख्य भाग में:

घरेलू मॉडल भारतीय भाषाओं और विशिष्ट जरूरतों के लिए क्षमता बढ़ाते हैं। स्थानीय इंफरेंस ढाँचा सेवा-उपलब्धता, संचालन-कौशल और उद्यमों की पहुँच बढ़ा सकता है। फिर भी विदेशी मॉडल को देश के सर्वर पर चलाने से प्रशिक्षण डेटा, लाइसेंस या महत्वपूर्ण तकनीक पर घरेलू नियंत्रण अपने-आप नहीं मिलता।

मूल्यांकन में डेटा-प्रवाह, लॉग रखने की अवधि, सुरक्षा-परीक्षण, लाइसेंस, कंप्यूटिंग उपलब्धता और प्रदाता बदलने की सुविधा शामिल हों। कंपनी का गति या लागत का दावा स्वतंत्र परीक्षण का स्थान नहीं लेता। सार्वजनिक खरीद में संवेदनशील डेटा तथा अपेक्षित जवाबदेही के अनुसार शर्तें तय हों।

निष्कर्ष में:

घरेलू नवाचार और वैश्विक तकनीक के विवेकपूर्ण उपयोग को साथ रखा जा सकता है। सफलता का पैमाना वास्तविक क्षमता, सुरक्षित उपयोग और जाँचने योग्य जवाबदेही हो।

PCSMANTRA.COM

अक्सर पूछे जाने वाले प्रश्न (FAQ)

क्या OpenAI-compatible का अर्थ OpenAI का मॉडल है?

नहीं। यहां यह API इंटरफेस के अनुरूप होने का दावा है, मॉडल के स्वामित्व या किसी दूसरी कंपनी के समर्थन का प्रमाण नहीं।

क्या स्थानीय होस्टिंग पूरी डेटा-सुरक्षा की गारंटी है?

नहीं। डेटा का संग्रह, लॉग, पहुँच-अधिकार, सेवा-शर्तें और सुरक्षा-परीक्षण भी देखने होते हैं।

क्या लॉन्च के प्रदर्शन-आंकड़े स्वतंत्र रैंकिंग हैं?

नहीं। वे कंपनी के बताए परीक्षण हैं। अलग परिस्थितियों के परिणामों को बिना समान परीक्षण-पद्धति के बराबर नहीं माना जा सकता।

स्रोत

1. https://www.sarvam.ai/blogs/sarvam-token-factory
2. https://www.pib.gov.in/PressReleasePage.aspx?PRID=2231169&lang=1&reg=3