Chariot-1: IndiaAI मिशन का नया स्वदेशी मॉडल

मूल प्रकाशन: 27 नवंबर 2026

समाचार-संदर्भ: PIB की 26 नवंबर 2026 की विज्ञप्ति के अनुसार IndiaAI मिशन के तहत चुनी गई कंपनी Chariot Artificial Intelligence Pvt. Ltd. के 'Chariot-1' फाउंडेशन मॉडल का शुभारंभ MeitY के सचिव श्री एस. कृष्णन ने किया।

प्रिलिम्स: करेंट अफेयर्स; विज्ञान-प्रौद्योगिकी (AI, भाषा-प्रौद्योगिकी), IndiaAI मिशन।
मुख्य परीक्षा: GS-3: विज्ञान एवं प्रौद्योगिकी में भारतीयों की उपलब्धियाँ, स्वदेशीकरण; GS-2: शासन में प्रौद्योगिकी।



pcsmantra.com

खबरों में क्यों

भारत अपने AI मॉडल खुद बनाने की दिशा में एक और कदम बढ़ा है। 26 नवंबर 2026 को जारी PIB विज्ञप्ति के अनुसार नई दिल्ली के भारत मंडपम में Chariot-1 फाउंडेशन मॉडल का शुभारंभ और शोध-प्रदर्शन हुआ। इसे Chariot Artificial Intelligence Pvt. Ltd. ने बनाया है। यह निजी कंपनी IndiaAI मिशन के तहत संप्रभु फाउंडेशन मॉडल विकसित करने के लिए चुनी गई पहली चार टीमों में शामिल है। मॉडल-शृंखला में एक भाषा मॉडल (LLM) और एक टेक्स्ट-टू-स्पीच मॉडल (TTS) है। दोनों भारत में शुरू से (from scratch) प्रशिक्षित बताए गए हैं।

बुनियादी बातें:
फाउंडेशन मॉडल: बहुत बड़े डेटा पर प्रशिक्षित ऐसा AI मॉडल, जिसे आगे अलग-अलग कामों के लिए ढाला जा सकता है।
पैरामीटर: मॉडल के भीतर सीखे हुए संख्यात्मक मान। पैरामीटर जितने ज़्यादा, मॉडल उतना बड़ा। बड़ा होने का मतलब अपने-आप बेहतर होना नहीं है।
मिक्सचर-ऑफ-एक्सपर्ट्स (MoE): ऐसी बनावट, जिसमें हर शब्द-इकाई (टोकन) पर पूरे मॉडल के बजाय उसका कुछ हिस्सा ही काम करता है। इससे गणना का खर्च घटता है।
टोकन: वे छोटे टुकड़े, जिनमें AI मॉडल पाठ को तोड़कर पढ़ता है।
टेक्स्ट-टू-स्पीच (TTS): लिखे पाठ को बोली गई आवाज़ में बदलने की तकनीक।

Chariot-1 क्या है

बनाने वाला: Chariot Artificial Intelligence Pvt. Ltd., एक निजी कंपनी। विज्ञप्ति में इसे IndiaAI मिशन के अंतर्गत विकसित बताया गया है।

भाषा मॉडल (LLM): 11 अरब पैरामीटर का मिक्सचर-ऑफ-एक्सपर्ट्स मॉडल, जिसमें प्रति टोकन 4 अरब पैरामीटर सक्रिय रहते हैं। यह तीन आकारों में उपलब्ध बताया गया है: 10.7 अरब, 1.7 अरब और 745 मिलियन पैरामीटर। विज्ञप्ति के अनुसार अलग-अलग आकार अलग-अलग कंप्यूटिंग माहौल के लिए हैं, यानी डेटा सेंटर से लेकर लैपटॉप, स्मार्टफोन और एज उपकरणों तक।

टेक्स्ट-टू-स्पीच (TTS): अंग्रेज़ी और 10 भारतीय भाषाओं में आवाज़ बनाता है: हिंदी, बंगाली, मराठी, तमिल, तेलुगु, कन्नड़, मलयालम, गुजराती, पंजाबी और ओड़िया। यानी यह संविधान की आठवीं अनुसूची की सभी 22 भाषाओं को कवर नहीं करता।

संख्याएँ, तिथियाँ, मुद्रा: विज्ञप्ति के अनुसार TTS मॉडल संदर्भ से समझता है कि संख्या, तारीख, मुद्रा, इकाई, URL और संक्षिप्त रूप कैसे बोले जाएँ।

तीन शोध नवाचार

विज्ञप्ति में Chariot टीम के तीन नवाचार गिनाए गए हैं, जो मॉडल के प्रशिक्षण में जोड़े गए। ये सभी आँकड़े डेवलपर-रिपोर्टेड हैं।

नवाचार	क्या करता है	बताया गया लाभ
IndicShift (ओपन-सोर्स)	भारतीय लिपियों के लिए साझा वर्णमाला बनाता है। 12 भारतीय लिपियों के 1,012 वर्णों को 350 अक्षरों से जोड़ता है।	OpenAI के GPT टोकनाइज़र की तुलना में भारतीय पाठ पर 2.5 गुना तक कम टोकन। समान कंप्यूट पर भारतीय भाषाओं में मॉडल-प्रदर्शन में 5% से 8% सुधार। अंग्रेज़ी और कोड में प्रदर्शन बरकरार।
Nested Training	एक ही प्रशिक्षण और एक ही वेट-सेट से तीन पूरे आकार के मॉडल साथ-साथ तैयार करता है।	तीनों आकारों को अलग-अलग प्रशिक्षित करने की तुलना में प्रशिक्षण की गति में 62% सुधार।
Mixture of Vocabularies	शब्दावली को 11 समूहों में बाँटकर एक समय पर केवल संबंधित समूह को जाँचता है।	प्रति शब्द जाँची जाने वाली प्रविष्टियाँ 5.7 गुना कम। उसी GPU पर प्रशिक्षण-गति में 38% सुधार, मॉडल-हानि में नगण्य बदलाव।

आँकड़ों को कैसे पढ़ें

TTS मॉडल के लिए विज्ञप्ति में 97% से अधिक 'वर्बलाइज़ेशन' सटीकता बताई गई है। यह Async TTS Pronunciation Benchmark पर मापी गई, जिसमें 1,029 वास्तविक वाक्य और 42 श्रेणियों के 2,251 गैर-मानक पाठ-उदाहरण हैं। हिंदी में कैरेक्टर एरर रेट 7.60% और अन्य नौ भारतीय भाषाओं में औसत 6.47% बताया गया है। विज्ञप्ति के अनुसार यह रिपोर्ट किए गए परीक्षणों में आँके गए तंत्रों में सबसे कम है।

सावधानी: ये नतीजे मॉडल बनाने वालों के बताए आँकड़े हैं और उन्हीं परीक्षणों तक सीमित हैं। इन्हें किसी स्वतंत्र रैंकिंग या सर्वश्रेष्ठ होने का प्रमाण न मानें। परीक्षा में इन्हें 'विज्ञप्ति के अनुसार रिपोर्ट किया गया' के रूप में ही याद रखना सुरक्षित है।

IndiaAI मिशन और संप्रभु मॉडल

विज्ञप्ति का तर्क है कि मिशन का उद्देश्य भारत की भाषाई, तकनीकी और सामाजिक ज़रूरतों के अनुरूप स्वदेशी AI क्षमता बनाना और बाहरी रूप से विकसित मॉडलों पर निर्भरता घटाना है। Chariot-1 इसी शृंखला की एक कड़ी है। इसी विषय पर pcsmantra पर पहले से उपलब्ध सामग्री भी देखें।

संबंधित लेख

सर्वम टोकन फैक्ट्री: AI आत्मनिर्भरता के अलग आयाम

IndiaAI कंप्यूट पोर्टल: सब्सिडी वाले GPU तक पहुँच

आगे की राह

स्वतंत्र मूल्यांकन: रिपोर्ट किए गए नतीजों की निष्पक्ष संस्थाओं या सार्वजनिक बेंचमार्क पर जाँच।
उपयोग: भाषा-सेवाओं, सरकारी सेवाओं और शिक्षा में वास्तविक तैनाती, ताकि मॉडल प्रयोगशाला से आगे जाए।
भाषाई विस्तार: TTS का दायरा 10 भारतीय भाषाओं से आगे बढ़ाना।
ओपन-सोर्स और डेटा: IndicShift जैसे ओपन-सोर्स घटकों को अपनाने और भारतीय भाषाओं के डेटा की गुणवत्ता पर काम।

निष्कर्ष

Chariot-1 IndiaAI मिशन के तहत एक निजी भारतीय कंपनी के बनाए संप्रभु मॉडल का उदाहरण है। परीक्षा की दृष्टि से तीन बातें याद रखें: मॉडल कौन-सा है (LLM और TTS), बनावट क्या है (MoE, 11 अरब में 4 अरब सक्रिय), और भाषाओं का दायरा (TTS में अंग्रेज़ी और 10 भारतीय भाषाएँ)।

प्रिलिम्स परीक्षा अभ्यास प्रश्न

Chariot-1 फाउंडेशन मॉडल के संदर्भ में निम्नलिखित कथनों पर विचार कीजिए:

कथन 1: इसे सरकारी संस्थान C-DAC ने विकसित किया है।

कथन 2: इसका भाषा मॉडल मिक्सचर-ऑफ-एक्सपर्ट्स बनावट का है, जिसमें प्रति टोकन केवल कुछ पैरामीटर सक्रिय रहते हैं।

कथन 3: इसका टेक्स्ट-टू-स्पीच मॉडल आठवीं अनुसूची की सभी 22 भाषाओं में आवाज़ बनाता है।

उपर्युक्त कथनों में से कौन-सा/से सही है/हैं?

(a) केवल 1 और 2

(b) केवल 2

(c) केवल 2 और 3

(d) 1, 2 और 3

▼ उत्तर और व्याख्या

उत्तर: (b)

कथन 1 गलत है: इसे निजी कंपनी Chariot Artificial Intelligence Pvt. Ltd. ने IndiaAI मिशन के तहत बनाया है।

कथन 2 सही है: 11 अरब पैरामीटर वाले मॉडल में प्रति टोकन 4 अरब सक्रिय रहते हैं।

कथन 3 गलत है: TTS अंग्रेज़ी और 10 भारतीय भाषाओं में आवाज़ बनाता है।

मुख्य परीक्षा अभ्यास प्रश्न

प्रश्न 1. भारत में संप्रभु AI फाउंडेशन मॉडल विकसित करने में IndiaAI मिशन की भूमिका और चुनौतियों की चर्चा कीजिए। (250 शब्द, 15 अंक)

उत्तर की रूपरेखा

भूमिका में: IndiaAI मिशन के तहत निजी टीमों द्वारा विकसित संप्रभु मॉडलों का संक्षिप्त संदर्भ; Chariot-1 का उदाहरण (LLM और TTS)।

मुख्य भाग में: भूमिका: स्वदेशी क्षमता, भाषाई विविधता के अनुरूप मॉडल (जैसे साझा लिपि-वर्णमाला, भारतीय भाषाओं में TTS), बाहरी निर्भरता में कमी, स्टार्टअप को सहारा। चुनौतियाँ: कंप्यूट और डेटा की उपलब्धता, भारतीय भाषाओं के उच्च-गुणवत्ता डेटा की कमी, डेवलपर-रिपोर्टेड नतीजों का स्वतंत्र सत्यापन, व्यावसायिक तैनाती और प्रतिभा को रोकना।

निष्कर्ष में: पारदर्शी बेंचमार्क, खुले घटकों और वास्तविक उपयोग पर ज़ोर देते हुए संतुलित निष्कर्ष।

अक्सर पूछे जाने वाले प्रश्न (FAQ)

संक्षिप्त उत्तर: Chariot-1 IndiaAI मिशन के तहत एक निजी कंपनी का बनाया स्वदेशी भाषा और आवाज़ मॉडल है। इसके प्रदर्शन के आँकड़े डेवलपर द्वारा रिपोर्ट किए गए हैं।

▼ प्रश्न 1. Chariot-1 को किसने बनाया?

Chariot Artificial Intelligence Pvt. Ltd. ने, जो IndiaAI मिशन के तहत संप्रभु मॉडल के लिए चुनी गई पहली चार टीमों में बताई गई है।

▼ प्रश्न 2. क्या Chariot-1 TTS सभी 22 भाषाओं में बोलता है?

नहीं। विज्ञप्ति के अनुसार यह अंग्रेज़ी और 10 भारतीय भाषाओं में आवाज़ बनाता है।

▼ प्रश्न 3. क्या 'सबसे कम त्रुटि दर' का मतलब स्वतंत्र रूप से सर्वश्रेष्ठ होना है?

नहीं। यह केवल रिपोर्ट किए गए परीक्षणों में आँके गए तंत्रों के बीच की तुलना है।

स्रोत

PIB: India AI Mission-supported Chariot-1 Foundation Models Launched (26 नवंबर 2026)

pcsmantra.com