Patrika (Indore), 30th September 2026, Page - 02

#AI: स्पीच ट्रांसलेशन में जेंडर की गलती सुधारने के लिए आइआइटी इंदौर ने बनाया एआइ टूल 'एचजीएएसटी'

अंग्रेजी में 'गया-गई' नहीं, एआइ ने निकाली बोलने वाले की पहचान

पत्रिका न्यूज नेटवर्क
patrika.com

इंदौर. 'आइ वेंट' अंग्रेजी में कोई पुरुष बोले या महिला, वाक्य एक ही रहेगा, लेकिन यही वाक्य जब हिंदी में बदलेगा तो पुरुष के लिए 'मैं गया' और महिला के लिए 'मैं गई' होगा। यही स्पीच ट्रांसलेशन करने वाले एआइ के सामने एक बड़ी चुनौती आती है। कई बार एआइ महिला के बोलने पर भी पुरुषवाचक शब्द का इस्तेमाल कर देता है। इस गलती को सुधारने के लिए आइआइटी इंदौर के शोधकर्ताओं ने हाइराकिंकल जेंडर आर्बिट्रेशन फॉर स्पीच ट्रांसलेशन (एचजीएएसटी) नाम का एआइ टूल विकसित किया है।

यह टूल आवाज सुनकर पहले यह पहचानता है कि बोलने वाला पुरुष है या महिला। इसके बाद अनुवाद में इस्तेमाल हुए क्रिया-रूप की जांच करता है और जरूरत पड़ने पर उसे सही करता है, यानी एआइ के अनुवाद के बाद एक तरह की 'जेंडर चेकिंग' होती है।

83 से बढ़कर 98.88% हुई महिला वक्ताओं की हिंदी में सटीकता



अंग्रेजी से हिंदी में ज्यादा दिखती है दिक्कत

अंग्रेजी में क्रियाओं का रूप वक्ता के जेंडर के हिसाब से नहीं बदलता, वहीं हिंदी, मराठी और गुजराती जैसी भारतीय भाषाओं में कई क्रियाओं के रूप पुरुष और महिला के अनुसार बदल जाते हैं। शोधकर्ताओं के अनुसार, एआइ अनुवाद प्रणालियों को ऐसे डेटा पर प्रशिक्षित किया जाता है, जिसमें पुरुषवाचक रूप अपेक्षाकृत अधिक हो सकते हैं। ऐसे में महिला की आवाज का अनुवाद करते समय भी सिस्टम पुरुषवाचक क्रिया चुन सकता है। इससे अनुवाद का अर्थ पूरी तरह नहीं बदलता।

पुराने सिस्टम को बदलने की जरूरत नहीं

एचजीएएसटी की खासियत यह है कि मौजूदा ट्रांसलेशन सिस्टम को दोबारा प्रशिक्षित करने की जरूरत नहीं है। यह पहले से काम कर रहे स्पीच ट्रांसलेशन सिस्टम के ऊपर 'करेक्शन लेयर' की तरह काम करता है। टूल पहले आवाज से जेंडर की जानकारी लेता है, फिर वाक्य की संरचना देखता है और यह जांचता है कि उसमें इस्तेमाल क्रिया सही जेंडर के अनुरूप है या नहीं। इसके बाद सत्यापित शब्दों के डेटाबेस और नियंत्रित एआइ भाषा मॉडल की मदद से सुधार करता है। सुधार के बाद दोबारा जांच भी की जाती है, ताकि वाक्य का अर्थ न बदले।

महिला वक्ताओं के अनुवाद में बड़ा सुधार

शोध टीम ने अंग्रेजी से हिंदी, मराठी व गुजराती में अनुवाद के दौरान एचजीएएसटी का परीक्षण किया। मौजूदा ट्रांसलेशन सिस्टम के साथ परीक्षण में महिला वक्ताओं के मामले में जेंडर की पहचान और सही क्रिया-रूप के इस्तेमाल में सुधार आया।

- **हिंदी:** 83.15% से बढ़कर 98.88%
- **गुजराती :** करीब 73-76% से बढ़कर करीब 96-97%
- **मराठी:** करीब 90-93% से बढ़कर 100%
- देशी मराठी वक्ताओं के स्वतंत्र परीक्षण में 98.2% सटीकता मिली
- पुरुष वक्ताओं के मामले में भी सिस्टम ने करीब 99-100% सटीकता बनाए रखी

एआइ की सामान्य रेटिंग से छिप जाती है ऐसी गलती

एआइ ट्रांसलेशन की गुणवत्ता जांचने के लिए इस्तेमाल होने वाले ब्लू और कोमेट जैसे सामान्य मानक जेंडर से जुड़ी छोटी गलतियों को पूरी तरह नहीं पकड़ पाते। शोधकर्ताओं के परीक्षण में कुछ मामलों में महिला वक्ताओं के लिए सुधार इन सामान्य मानकों में दिखाई देने वाले सुधार से 24 गुना से अधिक था। इसका मतलब यह है कि किसी अनुवाद में जेंडर से जुड़ी गलती हो सकती है।

500-500 लोगों ने की जांच

आइआइटी इंदौर के डॉ. चंद्रेश कुमार मौर्य के नेतृत्व और डॉ. महेंद्र गुप्ता के मार्गदर्शन में यह शोध किया गया। टीम में गवर्नमेंट कॉलेज ऑफ इंजीनियरिंग, अनूपपुर और इंटर्न अंजिल कुमार राज का भी सहयोग रहा। शोधकर्ताओं ने प्रत्येक भाषा के लिए 500 मानव समीक्षकों से एआइ के अनुवाद और सुधारों की मैन्युअल जांच कराई। मानव समीक्षा के परिणाम स्वचालित परीक्षणों से काफी हद तक मेल खाते पाए गए। शोध को यूरोपियन चैप्टर ऑफ द एसोसिएशन फॉर कम्प्यूटेशनल लिंग्विस्टिक्स के समक्ष एसीएल रोलिंग रिव्यू प्रक्रिया के माध्यम से पीयर रिव्यू के लिए प्रस्तुत किया गया है।