भारतीय भाषा आणि ज्ञान प्रणालीमध्ये AI

प्रा.कमलकिशोर उद्धवराव पंडागळे
सहाय्यक प्राध्यापक CHB, मराठी विभाग, श्रीमती वसल्याबाई नाईक महिला महाविद्यालय पुसद जि. यवतमाळ
Pandagalekk@gmail.com

प्रस्तावना

AI म्हणजे कृत्रिम बुद्धिमत्ता हा सद्यकाळात एक परवलीचा शब्द झाला आहे. सगळ्या काही गोष्टी एआयद्वारे साध्य होतात अशी काहीशी बहुतांश लोकांची समजूत झाली आहे. भाषाशास्त्राचा विचार करायचा झाला तर कुठलीही भाषा ही मानवाच्या संवादाचं मूळ साधन असते. निसर्गाने इतर प्राण्यांच्या तुलनेत मानवाला संवादासाठी भरभरून दिलेलं आहे आणि त्यामुळे हजारो वर्षांमध्ये मानवाच्या बुद्धीच्या विकासानुरूप भाषाही विकसित झाल्या आहेत. मूळ भाषा, बोलीभाषा, लिखित भाषा, व्याकरणशास्त्र, म्हणी, वाक्प्रचार, अलंकार असा सर्वच भाषांचा प्रचंड विस्तार आहे. भाषा ही मुळात संवादासाठी उत्पन्न झाली असली तरी कालौघात ती नुसती संवादाकरता न उरता पुस्तकांच्या माध्यमातून ज्ञानाची मोठ्या कालखंडासाठी साठवणूक आणि पर्यायाने ज्ञानार्जनासाठी अत्यंत महत्वाची झाली आहे.हे सर्व संगणकाला समजावून सांगणं आणि त्यात एआय म्हणजे कृत्रिम बुद्धिमत्ता वापरून भाषांतरं करणं आणि मानवी बुद्धीसारखं काम त्यांच्याकडून करून घेणं हे शिवधनुष्य पेलणं सोपं नाही. गूगल, मायक्रोसॉफ्टसारख्या कंपन्यांनी हे काम जागतिक भाषांसाठी हाती घेतलं आहे आणि थोड्या प्रमाणात यशस्वीही केलं आहे. पण अजूनही ही भाषांतरं यांत्रिक वाटतात. संदर्भानुसार शब्दयोजना, योग्य व्याकरण आणि वाक्यरचनेत यंत्रांना अजून खूप प्रगती करायची आहे. म्हणी, वाक्प्रचार, अलंकार आणि काव्य ही तर खुप दूरची गोष्ट आहे.

उद्देश

1) AI द्वारा भाषा विकास ही समजणे
2) भाषा शिक्षण मध्ये AI वापर करणे
3) AI भाषावर पडणारी सामाजिक व सांस्कृतिक विश्लेषण करणे

AI म्हणजे काय?

कृत्रिम बुद्धिमत्ता (एआय) म्हणजे यंत्रांद्वारे, विशेषतः संगणक प्रणालींद्वारे मानवी बुद्धिमत्ता प्रक्रियांचे अनुकरण. त्यात शिक्षण, तर्क आणि स्व-सुधारणा यांचा समावेश आहे. एआय अनुप्रयोगांच्या उदाहरणांमध्ये तज्ञ प्रणाली , नैसर्गिक भाषा प्रक्रिया (एनएलपी), भाषण ओळख , मशीन व्हिजन आणि चॅटजीपीटी आणि पर्प्लेक्सिटी सारखी जनरेटिव्ह टूल्स समाविष्ट आहेत.

कृत्रिम बुद्धिमत्ता किंवा AI कसं काम करते

मायकल गेरलिक स्वित्झर्लंडच्या एसबीएस बिझनेस स्कूलमध्ये मॅनेजमेंटचे प्राध्यापक आहेत. आम्ही त्यांना कृत्रिम बुद्धिमत्ता म्हणजे आर्टिफिशियल इंटेलिजंसची व्याख्या विचारली. ते माहिती देतात, "या संकल्पनेत दोन भाग आहेत — 'कृत्रिम' म्हणजे नैसर्गिक नसलेली किंवा मशीनद्वारा नियंत्रित केली जाणारी आणि 'बुद्धिमत्ता' म्हणजे विचार करण्याची क्षमता.

"कृत्रिम बुद्धिमत्ता म्हणजे AI माणसाच्या मेंदूच्या कार्यपद्धतीचीचं नक्कल करते. आपण मुलांना शिकवतो, तसंच AI ला प्रशिक्षण दिले जाते.

"त्यामुळे AI माणसासारखी तर्कशक्ती आणि समज विकसित करते. म्हणजेच उपलब्ध माहितीचे विश्लेषण करतं आणि अनेक कामांमध्ये आपली मदत करतं — इतकेच नव्हे तर काही वेळा आपल्यासाठी निर्णयही घेतं."

कृत्रिम बुद्धिमत्ता मोठ्या प्रमाणात माहितीचं वेगानं विश्लेषण करून उत्तरं शोधायला मदत कते.

भाषा प्रक्रिया ही कृत्रिम बुद्धिमत्तेची एक उपक्षेत्र म्हणून परिभाषित केली जाते जी मानवी भाषिक क्षमतांची प्रतिकृती तयार करण्यासाठी संगणकीय मॉडेल तयार करण्यावर लक्ष केंद्रित करते . यात भाषणाचा भाग टॅगिंग, नामांकित घटक ओळख आणि अर्थपूर्ण भूमिका लेबलिंग यासारख्या कार्यांचा समावेश आहे.

भाषा आणि AI प्रमुख पैलू

नैसर्गिक भाषा प्रक्रिया (एनएलपी): एआयचे हे क्षेत्र संगणकांना मानवी भाषा समजून घेण्यास, विश्लेषण करण्यास आणि प्रतिसाद देण्यास सक्षम करते.

नैसर्गिक भाषा प्रक्रिया

नैसर्गिक भाषा प्रक्रिया (NLP) मध्ये संगणकांना मानवांप्रमाणेच लिखित आणि बोलल्या जाणाऱ्या भाषा समजून घेण्यास आणि तयार करण्यास शिकवणे समाविष्ट आहे. NLP संगणक विज्ञान, भाषाशास्त्र, मशीन लर्निंग आणि सखोल शिक्षण संकल्पना एकत्रित करते जेणेकरून संगणकांना असंरचित मजकूर किंवा व्हॉइस डेटाचे विश्लेषण करण्यास आणि त्यातून संबंधित माहिती काढण्यास मदत होते. NLP प्रामुख्याने उच्चार ओळख आणि नैसर्गिक भाषा निर्मिती हाताळते आणि स्पॅम शोध आणि व्हर्च्युअल असिस्टंट सारख्या वापराच्या प्रकरणांमध्ये त्याचा वापर केला जातो .

अनुप्रयोग:

सामग्री निर्मिती : एआय स्वयंचलितपणे मजकूर आणि सामग्री तयार करू शकते.

भाषा भाषांतर : एआय वेगवेगळ्या भाषांमधील भाषांतर स्वयंचलित करते.

व्हॉइस असिस्टंट : एआय व्हॉइस कमांड समजू शकते आणि त्यांना प्रतिसाद देऊ शकते.

AI भाषा मॉडेल्स हे नैसर्गिक भाषा प्रक्रिया (एनएलपी) चा एक प्रमुख घटक आहेत, जे कृत्रिम बुद्धिमत्तेचे (एआय) क्षेत्र आहे जे संगणकांना मानवी भाषा समजून घेण्यास आणि निर्माण करण्यास सक्षम करण्यावर लक्ष केंद्रित करते.

AI भाषावर सामाजिक व सांस्कृतिक प्रभाव

1) कृत्रिम बुद्धिमत्ता कृत्रिम बुद्धिमत्ता AI आज मानव जीवनात प्रत्येक क्षेत्रात प्रवेश केला आहे आणि भाषा जे आपले विचार सांस्कृतिक परिचय करून देणे हे कृत्रिम बुद्धिमत्ता सामाजिक व सांस्कृतिक दोन्हीमध्ये महत्त्वाची भूमिका आहे

2) सामाजिक प्रभाव

अ) भाषेमध्ये असमानता प्रमुख भाषेत जसे इंग्रजी ,हिंदी, मराठी ,या अधिक विकसित होत आहे तसेच स्थानिक भाषाचे महत्व वाढले असून सामाजिक असमानता दिसून येत आहे

ब) भाषा शिकताना नवनवीन बदल होत आहे या मध्ये गुगल ट्रान्सलेशन मध्ये भाषा शिकणे अधिक सोपे झाले आहे. जसे की सांस्कृतिक सामाजिक व्यवहारात अधिक महत्व आले आहे .पारंपारिक व्यवहारात अधिक महत्व आपणास दिसून येईल.

3) सांस्कृतिक प्रभाव

भाषा डिजिटल करण होत आहे. यायचा उपयोग डेटाबेस असेल डिजिटल डेटाबेस असेल, यामध्ये भाषेत चे स्वरूप बदल होताना दिसते सांस्कृतिक दिशा सकारात्मक होताना दिसतो सांस्कृतिक भाषा ही आत्मा असते AI नुसार इंग्रजी भाषा प्राथमिकता देताना दिसते मातृभाषा पासून दूर जात आहे

AI मॉडेल्स आणि भाषा तंत्रज्ञान

सरकार भारताच्या स्वतःच्या पायाभूत मॉडेल्सच्या विकासास मदत करत आहे, ज्यामध्ये मोठ्या भाषा मॉडेल्स (LLMs) आणि भारतीय गरजांनुसार समस्या-विशिष्ट AI उपायांचा समावेश आहे. AI संशोधनाला चालना देण्यासाठी, अनेक उत्कृष्टता केंद्रे देखील स्थापन करण्यात आली आहेत.

भारताचे मूलभूत मोठ्या भाषेचे मॉडेल: इंडियाएआयने प्रस्ताव मागवून एलएलएम आणि स्मॉल लँग्वेज मॉडेल्स (एसएलएम) यासह स्वदेशी मूलभूत एआय मॉडेल्स विकसित करण्यासाठी एक उपक्रम सुरू केला आहे.

डिजिटल इंडिया भाषिनी: एक एआय-नेतृत्वाखालील भाषा भाषांतर प्लॅटफॉर्म जो भारतीय भाषांमध्ये इंटरनेट आणि डिजिटल सेवांमध्ये सहज प्रवेश प्रदान करण्यासाठी डिझाइन केलेला आहे, ज्यामध्ये व्हॉइस-आधारित प्रवेश आणि भारतीय भाषांमध्ये सामग्री निर्मितीला समर्थन देणे समाविष्ट आहे.

भारतजेन : जगातील पहिला सरकारी निधी असलेला मल्टीमॉडल एलएलएम उपक्रम, भारतजेन २०२४ मध्ये दिल्ली येथे सुरू करण्यात आला. भाषा, भाषण आणि संगणक दृष्टी या क्षेत्रातील मूलभूत मॉडेल्सद्वारे सार्वजनिक सेवा वितरण आणि नागरिकांचा सहभाग वाढवणे हे त्याचे उद्दिष्ट आहे. भारतजेनमध्ये भारतातील प्रमुख शैक्षणिक संस्थांमधील एआय संशोधकांचा एक संघ समाविष्ट आहे.

सर्वम-१ एआय मॉडेल : भारतीय भाषांसाठी अनुकूलित केलेले एक मोठे भाषा मॉडेल, सर्वम-१ मध्ये २ अब्ज पॅरामीटर्स आहेत आणि ते दहा प्रमुख भारतीय भाषांना समर्थन देते. हे भाषा भाषांतर, मजकूर सारांश आणि सामग्री निर्मिती यासारख्या अनुप्रयोगांसाठी डिझाइन केलेले आहे.

चित्रलेखा : AI4Bharat द्वारे विकसित केलेले एक ओपन-सोर्स व्हिडिओ ट्रान्सक्रिप्शन प्लॅटफॉर्म, चित्रलेखा वापरकर्त्यांना विविध भारतीय भाषांमध्ये ऑडिओ ट्रान्सक्रिप्ट तयार आणि संपादित करण्यास सक्षम करते.

हनुमानचा एव्हरेस्ट १.० : एसएमएलने विकसित केलेली बहुभाषिक एआय प्रणाली, एव्हरेस्ट १.० ३५ भारतीय भाषांना समर्थन देते, ज्याची संख्या ९० पर्यंत वाढवण्याची योजना आहे.

निष्कर्ष

AI नवीन भाषा व विकास करण्याचा प्रयत्न केला आहे. AI मुळे भाषेचा विकास अतिगतिमान व समृद्ध होत असल्याने अतिशय महत्त्व भाषा प्रणाली मध्ये महत्त्व प्राप्त झाले आहे. अनेक भाषांचे भाषांतर करून कमी कालावधीमध्ये ज्ञान प्राप्त करण्याचं एक महत्वपूर्ण काम AI मुळे होत आहे. AI मुळे भाषेची प्रगती होईल व सांस्कृतिक जडणघडण ला महत्त्व येऊ शकेल. भाषेबरोबर सांस्कृतिक जडणघडणीला महत्त्व प्राप्त होईल.

संदर्भ.

1. https://www.pib.gov.in/PressReleasePage.aspx?PRID=2108810
2. https://unique-features.com/articles/lekh/ai-for-bharat-bharatiy-bhashana-jodnara-prakalpa
3. https://www.fidelsoftech.com/news-and-blogs/ai-in-bhashinis-language-processing
4. https://www.deeplearning.ai/resources/natural-language-processing/