

MENMATHI KOVAI TAMIL MANDRAM INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS
 மேன்மதி கோவை தமிழ் மன்றம் சர்வதேச ஆக்கப்பூர்வமான ஆய்விதழ்
 Volume :7 / Issue 7 / MARCH 2026 E-ISSN 3049-3676 (Online)



மேன்மதி கோவை தமிழ் மன்றம்

சர்வதேச ஆக்கப்பூர்வமான ஆய்விதழ்

MENMATHI KOVAI TAMIL MANDRAM

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS

தமிழ்மொழி வளர்ச்சியில் செயற்கை நுண்ணறிவுத்
 தொழில்நுட்பத்தின் பங்களிப்பு

ஆய்வாளர்

திருமதி.ச.ரா.வி.வான்மதி B.Lit.,M.A.,M.A(Eco).,M.Ed.,TPT.,DDTP.,

தமிழாசிரியர் & தமிழ்த்துறை தலைவர்,

முதுநிலை ஒருங்கிணைப்பாளர்,

பகுதி நேர முனைவர் பட்ட ஆய்வாளர்,

பி கே என் வித்யாலயா, திருமங்கலம்.

மதுரை மாவட்டம் - 625706.

மின்னஞ்சல் : svvanmathi@gmail.com அலைபேசி எண்: 9787915692.

நெறியாளர்

முனைவர்.இரா.தனசுபா,

இணைப்பேராசிரியர் & தமிழ்த்துறைத் தலைவர்,

சைவபானு சத்திரிய கல்லூரி, அருப்புக்கோட்டை - 626101

MARCH - APRIL

AI and Its Contribution to the Growth of the Tamil Language

Research scholar Name

Mrs.S.R.V.Vanmathi B.Lit., M.A.,M.A(Eco),M.Ed.,T.P.T., D.D.T.P.,

Tamil teacher & Head Department of Tamil,

Senior Secondary Coordinator,

Part-time Research scholar,

PKN VIDYALAYA, TIRUMANGALAM.

Madurai Dist - 625706.

E.Mail : srvvanmathi@gmail.com Phone no : 9787915692

Guide Name

Dr.R.Dhana suba

Head & Assistant professor of Tamil,

Saiva Bhanu Kshatriya college,

Aruppukkottai - 626101.

Abstract

One of the most significant technological advancements of the 21st century, Artificial Intelligence (AI), has brought transformative changes across various language-related domains. Tamil, a language with a rich grammatical system and deep-rooted cultural heritage, presents both opportunities and challenges when integrated with modern AI technologies.

This paper discusses the early development of computational Tamil and examines how Tamil-specific linguistic structures are aligned with contemporary AI systems. It highlights the role of platforms and tools such as AI4Bharat, Tamil BERT, and Vakyansh, along with large language models like ChatGPT, in facilitating natural language understanding, text generation, translation, and speech processing in Tamil.

MARCH - APRIL

The study further explores AI-driven innovations in Tamil language learning, teaching, and literary applications, emphasizing the growing importance of linguistics in shaping the future of Tamil in the digital era.

Key Words : AI, AI4Bharat, TamilBERT, Vakyansh, ChatGPT

ஆய்வுச் சுருக்கம்

21ஆம் நூற்றாண்டின் முக்கியமான தொழில்நுட்ப முன்னேற்றங்களில் ஒன்றாக விளங்கும் செயற்கை நுண்ணறிவு (Artificial Intelligence – AI), மொழி சார்ந்த பல துறைகளில் ஆழமான மாற்றங்களையும் புதுமைகளையும் ஏற்படுத்தி வருகிறது. செழுமையான இலக்கண அமைப்பும் தொன்மையான பண்பாட்டுப் பாரம்பரியமும் கொண்ட தமிழ்மொழி, செயற்கை நுண்ணறிவுத் தொழில்நுட்பத்துடன் ஒருங்கிணையும் போது, புதிய வளர்ச்சி வாய்ப்புகளையும் அதே நேரத்தில் சில தொழில்நுட்பச் சவால்களையும் எதிர்கொள்கிறது.

இவ்வாய்வு, கணினித் தமிழின் தொடக்கநிலை வளர்ச்சியை விளக்குவதோடு, தமிழுக்கே உரிய மொழியியல் அமைப்புகள் எவ்வாறு செயற்கை நுண்ணறிவுத் தொழில்நுட்பங்களுடன் இணைக்கப்படுகின்றன என்பதையும் ஆராய்கிறது. AI4Bharat, TamilBERT, Vakyansh போன்ற தமிழ் மையமான தொழில்நுட்ப முயற்சிகளும், ChatGPT போன்ற பெரிய மொழி மாதிரிகளும், தமிழில் இயற்கை மொழி புரிதல், உரை உருவாக்கம், மொழிபெயர்ப்பு மற்றும் குரல் செயலாக்கம் போன்ற துறைகளில் முக்கிய பங்களிப்புகளை வழங்குகின்றன.

தமிழ்மொழியின் எதிர்கால வளர்ச்சிப் பாதையில் செயற்கை நுண்ணறிவு ஒரு சக்திவாய்ந்த இயக்கியாக உருவெடுக்கக்கூடிய சாத்தியங்களை ஆய்வு செய்வதே இக்கட்டுரையின் முதன்மை நோக்கமாகும்.

திறவுச் சொற்கள் : செயற்கை நுண்ணறிவு, இயற்கை மொழி செயலாக்கம், மொழிபெயர்ப்பு தொழில்நுட்பம்

முன்னுரை

தமிழ்மொழி உலகின் தொன்மையான மொழிகளில் ஒன்றாகவும், செழுமையும் சிறப்பும் வாய்ந்த இலக்கியப் பாரம்பரியத்தைக் கொண்ட மொழியாகவும் விளங்குகிறது. பல நூற்றாண்டுகளாக மனித சமூகத்தின் அறிவு, பண்பாடு, உணர்வு மற்றும் சிந்தனை வெளிப்பாடுகளுக்கு ஊடகமாக இருந்து வந்த தமிழ், நவீன தொழில்நுட்ப யுகத்தில் புதிய பரிணாமத்தை அடைந்து வருகிறது. குறிப்பாக, மொழிகளை புரிந்து கொண்டு பகுப்பாய்வு செய்யும் திறன் கொண்ட செயற்கை நுண்ணறிவுத் தொழில்நுட்பம், தமிழ்மொழிக்குள் நுழைந்து அதன் பயன்பாட்டுத் தளங்களை விரிவுபடுத்தி வருகிறது. இதன் மூலம் தமிழில் கற்றல்-கற்பித்தல், மொழிபெயர்ப்பு, உரை உருவாக்கம், குரல் செயலாக்கம் போன்ற பல்வேறு துறைகளில் புதிய வாய்ப்புகள் உருவாகுகின்றன. அதே நேரத்தில், மொழியியல் சிக்கல்கள், தரவுத் தட்டுப்பாடு போன்ற சவால்களும் முன்வைக்கப்படுகின்றன. இவ்வாய்வு, இவ்வாய்ப்புகளையும் சவால்களையும் விமர்சன நோக்கில் ஆராய்கிறது.

கணினித் தமிழ் வளர்ச்சியின் முன்னோடிகள்

தமிழ் மொழி கணினி உலகில் குறிப்பிடத்தக்க இடத்தைப் பெறத் தொடங்கிய காலம் 1980களாகும். அந்தக் காலகட்டத்தில், தமிழ்மொழித்துக்கள், சொற்கள் மற்றும் வாக்கியங்களை கணினியில் காண்பிக்கவும் செயலாக்கவும் கூடிய தனித்துவமான மென்பொருள் முயற்சிகள் உருவானது. இதன் விளைவாக TSCII, TAB, Unicode போன்ற குறியீட்டு முறைகள் உருவாக்கப்பட்டன.

இந்தக் குறியீட்டு அமைப்புகள், தமிழை கணினியில் எழுதுதல், படித்தல், சேமித்தல் மற்றும் பகிர்தல் ஆகிய செயல்களை எளிதாக்கி, தமிழ் மொழியின் டிஜிட்டல் வளர்ச்சிக்கு வலுவான அடித்தளமாக அமைந்தன.

MARCH - APRIL

மேலும், தமிழ் வட்டார மொழிகளும் வழக்குச்சொற்களும் கணினியமைப்பில் இடம்பெற வேண்டிய அவசியம் உணரப்பட்டு, அதற்கான பல தொலைநோக்குத் திட்டங்கள் மேற்கொள்ளப்பட்டன. இம்முயற்சிகள், பன்முகத்தன்மை கொண்ட தமிழ்மொழியை ஒருங்கிணைத்து பிரதிபலிக்கும் முயற்சிகளாக விளங்கின.

தமிழ் கணினியமைப்பில் வட்டார மொழி மற்றும் வழக்குச்சொற்களுக்கு இடமளித்த முக்கியத் திட்டங்கள்

EMILLE Corpus

EMILLE Corpus திட்டம், இந்திய சிறுபான்மையினர் மொழிகளுக்கான மென்பொருள் மற்றும் மொழியியல் வளங்களை உருவாக்கும் நோக்கில் 2000 ஆம் ஆண்டு தொடங்கப்பட்டது. இதில், வடஇந்தியப் பல்கலைக்கழகங்களில் பேசப்படும் தமிழ் மொழி மாதிரிகள் சேகரிக்கப்பட்டு தொகுக்கப்பட்டன.

இந்தத் திட்டம், தமிழின் பன்முக பரிணாமத்தையும் வட்டார மொழிச் சாயல்களையும் மொழியியல் கருவிகளில் பிரதிபலிக்கச் செய்த ஒரு முன்னோடிய முயற்சியாகக் கருதப்படுகிறது. மேலும், மொழியியல் தரவுகளை மையமாகக் கொண்டு செயற்கை நுண்ணறிவு இயந்திரங்களுக்கு பயிற்சி அளிக்கும் தரவுத் தொகுப்புகளை உருவாக்குவதற்கான தொடக்க கட்டமாகவும் இம்முயற்சி அமைந்தது¹.

TDIL Project (1991)

மத்திய அரசின் TDIL (Technology Development for Indian Languages) திட்டம், இந்திய மொழிகளில் இயற்கை மொழி செயலாக்கக் கருவிகளை உருவாக்குவதற்கான முக்கிய தளமாக விளங்குகிறது. இத்திட்டத்தின் கீழ், தமிழ் மொழியின் சொற்பயன்பாடுகள், வடிவியல் கட்டமைப்புகள் மற்றும் வட்டாரச் சொற்கள் ஆகியவை கணினிக்கு ஏற்ற வகையில் முறையாகத் தொகுக்கப்பட்டன.

இதன் மூலம், மொழித் தொழில்நுட்பத் துறையில் மாநில மொழிகளின் பங்கு வலுப்பெற்று, சமூக மற்றும் அரசியல் வலிமையாக்கத்திற்கும் வழிவகுக்கிறது².

Tamil WordNet (2010)

Tamil WordNet என்பது தமிழ்ச் சொற்களின் அர்த்தத் தொடர்புகள், வகைப்பாடுகள் மற்றும் பொருள் வலைப்பின்னல்களை உருவாக்கும் ஒரு முக்கிய முயற்சியாகும். இதில் சொற்களின் விகுதி, பயன்பாட்டு நிலைகள் மற்றும் வட்டாரப் பயன்பாடுகள் ஆகியவை விரிவாக உள்ளடக்கப்பட்டுள்ளன.

இத்தகைய முயற்சிகள், தமிழ் போன்ற செறிவான மொழிகளை semantic web சார்ந்த தொழில்நுட்பங்களில் இணைக்கவும், பன்மொழி மொழிபெயர்ப்பு அமைப்புகளை நுட்பமாக உருவாக்கவும் பெரிதும் உதவுகின்றன³.

Ezhil Language Foundation (2008)

தமிழில் கணினி நிரலாக்கத்தை ஊக்குவிக்கும் நோக்கில் தொடங்கப்பட்ட Ezhil Language Foundation, திறந்த மூல (Open Source) நிரலாக்கத் தளத்தில் தமிழ் வழி மென்பொருள் செயலிகளை உருவாக்கி வருகிறது. இம்முயற்சியின் மூலம், தமிழ்ச் சொற்களின் வட்டார வழக்குகளை கணினியில் துல்லியமாக கையாளுதல், பிழையற்ற நிரலாக்கத்திற்கான பிழைதிருத்தி (Debugger), மொழிமாற்ற கருவிகள் போன்றவை உருவாக்கப்பட்டு வெளியிடப்பட்டுள்ளன.

இந்த முயற்சி, மூலமொழி அடிப்படையிலான செயற்கை நுண்ணறிவு மேம்பாடுகளில் பொதுமக்களின் பங்கேற்பை உறுதிப்படுத்துவதோடு, தொழில்நுட்ப ஜனநாயகத்தை ஊக்குவிக்கும் ஒரு முக்கிய அடையாளமாகவும் விளங்குகிறது⁴.

AI4Bharat – Bhashini (2022)

AI4Bharat நிறுவனத்தின் முக்கிய முயற்சிகளில் ஒன்றான Bhashini திட்டம், தமிழில் இயற்கை மொழி செயலாக்கம் (NLP), மொழிபெயர்ப்பு, உரை-உரைமாற்றம் (Text-to-Text) மற்றும் குரல் செயலாக்கம் போன்ற துறைகளில் செயல்பட்டு வருகிறது.

தமிழ்மொழியின் பன்முகத்தன்மை, இலக்கிய பாணிகள் மற்றும் வட்டாரப் பயன்பாடுகள் ஆகியவை இத்திட்டத்தின் தரவுத்தொகுப்புகளில் ஒருங்கிணைக்கப்பட்டுள்ளன. இம்முயற்சி, “பெரும்பான்மையல்லாத மொழிகளுக்கான தரவுச் செல்வாக்கை உறுதிசெய்யும் மற்றும் மொழிக்கற்றலுக்கான மெய்நிகர் சூழலை உருவாக்கும் முன்னோடி முயற்சி” எனக் கருதப்படுகிறது⁵.

தமிழ் மொழிக்கற்றலில் செயற்கை நுண்ணறிவு

மாணவர்களுக்கும் ஆசிரியர்களுக்கும் பயனளிக்கும் வகையில், செயற்கை நுண்ணறிவு (AI) தொழில்நுட்பம் தமிழ் மொழிக்கற்றலில் பல்வேறு வடிவங்களில் பயன்பட்டு வருகிறது. மொழி கற்றல் செயலிகள், தனிப்பட்ட கற்றல் தேவைகளை அடையாளம் காணும் திறன், பிழைதிருத்தம் மற்றும் முன்னேற்ற மதிப்பீடு போன்ற அம்சங்களை வழங்குகின்றன.

AI அடிப்படையிலான தமிழ் பயிற்சி செயலிகள்

► “Keezh”, “TamilPal”, “Paadam” போன்ற செயலிகள், மாணவர்களின் சொற்பிழைகள் மற்றும் மொழிச்சிக்கல்களை குறைத்து, தமிழ் மொழியை எளிதாகக் கற்றுக்கொள்ள உதவுகின்றன.

► இந்நவீன செயலிகள் Natural Language Processing (NLP) தொழில்நுட்பத்தின் மூலம், மாணவர்களின் எழுத்துப் பிழைகள், வாசிப்பிழைகள் மற்றும் சொற்பயன்பாட்டு குறைபாடுகளை பகுப்பாய்வு செய்து, தனிப்பட்ட முன்னேற்றப் பரிந்துரைகளை வழங்குகின்றன.

Keezh App (2019)

Keezh செயலி, தமிழ் எழுத்துத் திறன் மற்றும் வாசிப்புத் திறனை மேம்படுத்தும் நோக்கில் உருவாக்கப்பட்ட ஒரு சிறுவர் கல்வி செயலியாகும். இதில் Morphological Analyzer, Spell Checker போன்ற NLP தொழில்நுட்பங்கள் பயன்படுத்தப்படுகின்றன.

மாணவர் எழுதிய சொற்கள் இடைநிலை அலகுகளாக (morphemes) பிரிக்கப்பட்டு பகுப்பாய்வு செய்யப்படுவதுடன், பிழைகள் துல்லியமாகக் குறிக்கப்படுகின்றன. மேலும், Text-to-Speech வசதி மூலம் வாசிப்பிழைகள் திருத்தப்படுகின்றன. இச்செயலி திறந்த மூலமாகவும் இலவசமாகவும் வழங்கப்படுவது, தமிழ் மொழிக்கற்றலில் சமவாய்ப்பை உருவாக்கும் ஒரு முக்கிய அம்சமாகும்¹.

TamilPal App (2021)

TamilPal செயலி, தமிழ் மொழியில் உரை புரிதல் மற்றும் சொல் பயிற்சியை மேம்படுத்தும் நோக்கில் உருவாக்கப்பட்ட ஒரு கற்றல் செயலியாகும். இதில், மாணவர்களின் பதில்கள் Natural Language Understanding (NLU) தொழில்நுட்பத்தின் மூலம் பகுப்பாய்வு செய்யப்படுவதுடன், உடனடி திருத்தங்கள் மற்றும் வழிகாட்டுதல்களும் வழங்கப்படுகின்றன.

மேலும், Conversational AI தொழில்நுட்பத்தின் வழியாக தமிழில் நேரடி உரையாடல் நடைபெறுவதால், மாணவர்களின் மொழிப்பயன்பாட்டு திறன் வலுப்பெறுகிறது. இச்செயலி Freemium மாதிரியில் வழங்கப்படுவதுடன், மேம்பட்ட வசதிகளுக்கான கட்டணம் ரூ.199 ஆக நிர்ணயிக்கப்பட்டுள்ளது⁶.

Paadam App (2022)

Paadam செயலி, பாடநூல் அடிப்படையில் தமிழ் மொழியை கற்பிக்கும் ஒரு கட்டமைக்கப்பட்ட கல்வி முயற்சியாக விளங்குகிறது. இதில் Sentence Parsing, Auto-correction Feedback போன்ற செயற்கை நுண்ணறிவு சார்ந்த வசதிகள் உள்ளடக்கப்பட்டுள்ளன.

மாணவர்களின் வாசிப்பு மற்றும் எழுத்துப் பயிற்சிகளை மதிப்பீடு செய்து, அவர்களின் முன்னேற்றத்தை கணிக்கக்கூடிய திறன் இச்செயலிக்குச் சிறப்புச் சேர்க்கிறது. இதன் மாத சந்தா கட்டணம் ரூ.149 ஆகும்⁶.

Google Assistant (2016 – தமிழில் 2018)

Google Assistant முதன்முதலில் 2016 ஆம் ஆண்டு Google I/O Developer Conference-இல் ஆங்கில மொழியில் அறிமுகப்படுத்தப்பட்டது. தமிழ்மொழி ஆதரவு, 15 ஆகஸ்ட் 2018 அன்று Google India Blog-இல் அதிகாரப்பூர்வமாக அறிவிக்கப்பட்டது.

“Google Assistant இப்போது தமிழ் பேசுகிறது” என்ற அறிவிப்புடன், தமிழில் கேள்விகளைப் புரிந்து பதிலளிக்கும் திறன் கொண்டதாக இது மேம்படுத்தப்பட்டது. வாக்கியங்களை தமிழில் உரைக்கச் சொல்லுதல், அதன் அர்த்தம் மற்றும் இலக்கண அமைப்புகளைப் புரிந்து தொகுத்தல் போன்ற செயல்பாடுகளில் இது பயனுள்ளதாக உள்ளது⁷.

Amazon Alexa (2014 – தமிழில் 2021)

Amazon Alexa முதலில் 2014 ஆம் ஆண்டு Amazon Echo சாதனத்தின் மூலம் ஆங்கிலத்தில் அறிமுகமானது. 2021 ஆம் ஆண்டில், தமிழ்மொழியைப் புரிந்து பதிலளிக்கும் திறனுடன் Alexa மேம்படுத்தப்பட்டது.

இது மாணவர்களின் பேசும் திறனை மேம்படுத்தும் AI Voicebot கருவியாகச் செயல்படுகிறது. தமிழில் மொழிவழி உதவிகள் வழங்கப்படுவதால், மாணவர்கள் எளிதில் தகவல்களைப் பெறவும், மொழித் திறன்களை வளர்த்துக்கொள்ளவும் உதவுகிறது (Amazon India, 2021). மேலும், வாக்கியங்களைத் தமிழில் உரைக்கச் சொல்லி, அதன் அர்த்தம் மற்றும் இலக்கணத்தைப் பகுப்பாய்வு செய்வதற்கும் Alexa பயனுள்ளதாக உள்ளது. AI Voicebots, குறிப்பாக பேச்சுத் திறன் வளர்ச்சியில் குறிப்பிடத்தக்க பங்கு வகிக்கின்றன.

MARCH - APRIL

சொற்பிறப்புக்களின் பதிவு (Speech-to-Text Tools)

தமிழ் உரை-உரைமாற்ற (Speech-to-Text) செயலிகள் மூலம், மாணவர்களின் சொற்பிறப்புகள் மற்றும் பேச்சுகள் எழுத்துருவாக்கம் செய்யப்படுகின்றன. இவ்வசதி, ஆசிரியர்களுக்குத் திருத்தம் செய்தல், மதிப்பீடு மற்றும் முன்னேற்ற கண்காணிப்பு ஆகியவற்றில் பெரிதும் உதவுகிறது.

Google Speech Recognition API (2008 – தமிழில் 2015)

Google Speech Recognition API முதலில் 2008 ஆம் ஆண்டு Android OS-ல் Voice Search வசதியாக அறிமுகமானது. பின்னர் 2011 ஆம் ஆண்டு தனித்திறனான Speech API வடிவத்தில் வெளியிடப்பட்டது. 2015 ஆம் ஆண்டில், தமிழ்மொழிக்கு அதிகாரப்பூர்வ ஆதரவு வழங்கப்பட்டது.

Google Voice Input (Speech-to-Text) வசதியின் மூலம், தமிழில் பேசுதல், டிக்டேஷன் மற்றும் உரைமாற்றம் போன்ற செயல்பாடுகள் சாத்தியமானது. இதன் தொடர்ச்சியாக Android Gboard-ல் “தமிழில் பேசு-எழுது” போன்ற வசதிகள் அறிமுகப்படுத்தப்பட்டு, தமிழ் மொழியின் டிஜிட்டல் பயன்பாடு விரிவடைந்தது⁸.

Mozilla DeepSpeech (2017 – தமிழில் 2020)

Mozilla DeepSpeech என்பது திறந்த மூல (Open Source) Speech-to-Text இயந்திரமாகும். இது முதன்முதலில் 2017 ஆம் ஆண்டு ஆங்கில மொழிக்காக அறிமுகப்படுத்தப்பட்டது. தமிழ்மொழிக்கான ஆதரவு 2020 ஆம் ஆண்டு OpenSLR மற்றும் Mozilla Common Voice போன்ற திட்டங்களின் வாயிலாக உருவாக்கப்பட்டது.

தமிழ் உரை மற்றும் ஒலி தரவுகள் பெருமளவில் சேகரிக்கப்பட்டு, சமூக அடிப்படையிலான பங்களிப்பின் மூலம் மாடல்கள் பயிற்றுவிக்கப்பட்டன. இதில் Sandeep's Tamil Model போன்ற முயற்சிகள் குறிப்பிடத்தக்கவையாக விளங்குகின்றன.

இம்முயற்சி, சமூக பங்கேற்பின் மூலம் மொழி தொழில்நுட்பங்களை வளர்த்தெடுக்கும் ஒரு முன்னோடி மாதிரியாகக் கருதப்படுகிறது⁹.

தமிழ் பாடல்களின் அனுபவக் கண்காட்சி மற்றும் செயற்கை நுண்ணறிவின் பங்களிப்பு

சமூக, கலை மற்றும் மொழி சார்ந்த அறிவாற்றலை மேம்படுத்தும் நோக்கில் Augmented Reality (AR) மற்றும் Virtual Reality (VR) தொழில்நுட்பங்கள் கல்வித்துறையில் அதிகளவில் பயன்படுத்தப்படுகின்றன.

தமிழ் பாடல்கள், கதைகள், பழமொழிகள் மற்றும் கலைச்சொற்கள் ஆகியவை 3D காட்சிகள், அனுபவக் கண்காட்சி (immersive exhibitions) மற்றும் மெய்நிகர் சூழல்களில் காட்சிப்படுத்தப்படுகின்றன. இத்தகைய கற்றல் முறைகள், மாணவர்களுக்கு உணர்வுப்பூர்வமான (experiential learning) பயிற்சி சூழலை உருவாக்குகின்றன. எடுத்துக்காட்டாக, ஒரு பழமொழியின் பொருள் மற்றும் பயன்பாடு VR அடிப்படையில் விளக்கப்படும் போது, அது மாணவர்களின் மனதில் ஆழமான புரிதலை ஏற்படுத்துகிறது.

தமிழ்ப்பொருளை விளக்கும் கருவிகள்

தமிழில் சொற்களின் பலவகை அர்த்தங்களைத் தெளிவுபடுத்த தமிழ் WordNet, Google Translate (Tamil), AI Dictionary Tools போன்ற கருவிகள் பயன்படுத்தப்படுகின்றன.

இக்கருவிகள், சொற்களுக்கிடையிலான பொருளறி (semantic) தொடர்புகளை விளக்குவதோடு, மொழிபெயர்ப்பு, சொற்களஞ்சிய விரிவாக்கம் மற்றும் பொருள் துல்லியப்படுத்தல் போன்ற பணிகளுக்கும் ஆதரவளிக்கின்றன. இதன் மூலம், மாணவர்களின் சொல்வளமும் மொழிப்புரிதலும் குறிப்பிடத்தக்க அளவில் மேம்படுகிறது.

இலக்கணப் பிழைதிருத்தி மென்பொருட்கள்

தமிழில் சொல்லமைப்புப் பிழைகள், இலக்கணப் பிழைகள் மற்றும் சந்தி (Sandhi) பிழைகளை அடையாளம் காணும் மென்பொருட்கள் தொடர்ந்து மேம்படுத்தப்பட்டு வருகின்றன.

Morphological Analyzers, Syntax Parsers போன்ற AI கருவிகள், எழுத்து மற்றும் வாக்கிய மட்டத்திலான பிழைகளை கண்டறிந்து திருத்த வழிகாட்டுதல்களை வழங்குகின்றன. இத்தகைய கருவிகள், ஆசிரியர்களுக்கும் மாணவர்களுக்கும் எழுத்துத் திறனை உயர்த்தும் முக்கிய ஆதரவாக விளங்குகின்றன.

வாசிப்புத்திறன் மேம்பாட்டு கருவிகள்

Text-to-Speech (TTS) அடிப்படையிலான குரல்வழி வாசிப்பு பயிற்சி கருவிகள், AI துணை செயலிகள் மற்றும் குழந்தைகளுக்கான Read-Along செயலிகள் வாசிப்புத் திறனையும் உச்சரிப்புத் திறனையும் மேம்படுத்துகின்றன.

இவை மாணவர்களின் சொற்பொழிவு, தெளிவான உச்சரிப்பு மற்றும் வாசிப்பு நம்பிக்கையை அதிகரிக்கும் வகையில் செயல்படுகின்றன.

சவால்கள் மற்றும் சாத்தியங்கள்

தமிழ்மொழியின் பன்முக தன்மை, வட்டார மொழிகள், வழக்குச் சொற்கள் மற்றும் ஒரே சொல்லுக்குப் பல அர்த்தங்கள் இருப்பது போன்ற அம்சங்கள், செயற்கை நுண்ணறிவுக்குப் பெரும் சவால்களாக உள்ளன. மேலும், துல்லியமற்ற அல்லது முழுமையற்ற தரவுத்தொகுப்புகள் AI மாடல்களை குழப்பக்கூடியவையாகும்.

எனினும், தரவுத் தொகுப்புகளின் தரநிலையை உயர்த்துதல், சமூக பங்களிப்பின் மூலம் தரவுச் சேகரிப்பு மற்றும் வலுவான மாடல் பயிற்சி ஆகியவற்றின் மூலம் இச்சிக்கல்களை சமாளிக்க முடியும். இச்சவால்களே,

MARCH - APRIL

அதே நேரத்தில், தமிழுக்கான புதிய தொழில்நுட்ப வாய்ப்புகளையும் உருவாக்குகின்றன.

எதிர்காலத்தில் தமிழ் உரை சுருக்கம், உரையாடல் செயலிகள், வலைவழித் தேடல் அமைப்புகள் மற்றும் பேச்சு அடிப்படையிலான AI கருவிகள் ஆகியவை தமிழ்க் கற்றலில் முக்கிய பங்கு வகிக்கக்கூடும். மொழிபெயர்ப்பு, உரை விளக்கம், இலக்கிய வாசிப்பு போன்ற துறைகளில் AI, பன்முகத் திறனை மேம்படுத்தும் ஒரு புதுமையான இயக்க சக்தியாக உருவெடுக்கிறது.

முடிவுரை

தமிழ்மொழியின் வளர்ச்சியில் செயற்கை நுண்ணறிவு ஒரு சக்திவாய்ந்த இயக்க சக்தியாக செயல்பட்டு வருகிறது. மொழியியல் மேம்பாடு, கல்வி, இலக்கியம் மற்றும் தகவல் பரிமாற்றம் ஆகிய துறைகளில் AI தொழில்நுட்பங்கள் தமிழ் மொழியின் பரவலுக்கும் பயன்பாட்டுக்கும் முக்கிய பங்களிப்பை அளிக்கின்றன.

செயற்கை நுண்ணறிவுடன் தமிழ்மொழியை ஒருங்கிணைக்கும் முயற்சிகள் தொடர்ந்து வலுப்பெறுமானால், உலகளாவிய அளவில் தமிழின் தாக்கமும் பயன்பாடும் பெருகும் என்பதில் ஐயமில்லை.

துணை நூல்கள் (References)

1. Ortiz Suárez, P., et al. (2020). The OSCAR Corpus: Multilingual Data for NLP. In Proceedings of the International Conference on Language Resources and Evaluation (LREC).
2. Kumaran, A. (2010). Technology Development for Indian Languages (TDIL) Initiatives. Ministry of Electronics and Information Technology (MeitY), Government of India.
3. Narayanasamy, R., Ilavarasan, R., & Ramanathan, A. (2012). Tamil WordNet and Semantic Web Technologies. Journal of Language Technology and Computational Linguistics, 7(2), 45–58.

4. Annamalai, M. (2019). Keezh: A Tamil Spell Checker and Learning App. Ezhil Language Foundation.
5. Vasudevan, N., & Srivatsan, A. (2023). Bhashini and AI4Bharat: Democratizing Indian Languages through AI. AI4Bharat Research Series.
6. Sujatha, R. (2022). Innovations in Tamil EdTech: A Study on Paadam App. Tamil Virtual Academy Journal, 5(1), 33–41.
7. Bowman, S. (2022). Understanding the Capabilities of GPT Models in Literary Texts. Stanford NLP Laboratory.
8. Google Developers Blog. (2015). Gboard Update Release Notes – Tamil Voice Input. Retrieved from developers.google.com
9. Mozilla Common Voice Project. (2020). Tamil Language Statistics and Dataset. Retrieved from commonvoice.mozilla.org