Tamilmanam – International Research Journal of Tamil Studies
An international, peer-reviewed academic journal
e-ISSN: 3049-0723
pISSN:3107-7781

தமிழ்மணம்
Peer-reviewed · Open access · Google Scholar · Crossref DOI · SJIF
Imapact Factor 3.5 · Monthly
Vol. 20, No. 02 (2026) 20 02 2026 155–161 TIRJTS-2026-0007

UGC Approved by Transparent Peer Review

RESEARCH ARTICLE

இயற்கை மொழி செயலாக்கம்

Natural Language Processing

Dr. M. Palaniyammal[1,*]

[1] Assistant Professor, Department of Tamil, Mohamed Sathak College of Arts and Science, Sholinganallur, Chennai – 600119

***Corresponding author:** Dr. M. Palaniyammal, palaniyammal.anand@gmail.com

ARTICLE INFO

Article history
Received: 1 October 2026

Article ID
TIRJTS-2026-0007

Peer review
Double-blind peer review

Keywords
Natural Language Processing
Artificial Intelligence
Computational Linguistics
Machine Translation
Tamil Language
Text Analysis
Language Technology.

ABSTRACT

Artificial Intelligence plays a significant role in today's scientific and technological development. Natural Language Processing (NLP) is one of the important branches of Artificial Intelligence. The fundamental objective of NLP is to enable computers to understand, analyse, interpret, and respond appropriately to human languages such as Tamil and English. NLP simplifies human–computer interaction and has applications in various fields, including machine translation, text-to-speech conversion, speech-to-text conversion, sentiment analysis, question-answering systems, information retrieval, and text summarisation.

Natural Language Processing offers new opportunities for languages with rich literary traditions, such as Tamil. However, the unique characteristics of the Tamil language, including its script, word structure, grammatical system, inflectional morphology, semantic variations, and dialectal differences, present several challenges in NLP research and development. These linguistic features require specialised computational approaches and language-specific resources to improve the accuracy and effectiveness of language-processing systems.

This paper examines the origin and development of Natural Language Processing, its fundamental components, major applications, and current status in Tamil language technology. It also discusses the challenges involved in processing Tamil through computational methods and explores future opportunities for developing effective NLP tools and applications for Tamil. The study highlights the importance of integrating linguistic knowledge with modern computational techniques to preserve, promote, and expand the use of the Tamil language in the digital era.

How to cite this article: Dr. M. Palaniyammal (2026). இயற்கை மொழி செயலாக்கம். *Tamilmanam – International Research Journal of Tamil Studies*, *20*(02), 155–160.
Copyright and licence: © 2026 The Author(s). This is an open access article distributed under the terms of the Creative

Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

ஆய்வுச் சுருக்கம்

இன்றைய அறிவியல் மற்றும் தொழில்நுட்ப வளர்ச்சியில் செயற்கை முக்கியப் பங்கு வகிக்கிறது. செயற்கை நுண்ணறிவின் குறிப்பிடத்தக்க கிளைகளுள் ஒன்றாக இயற்கை மொழி விளங்குகிறது. மனிதர்கள் இயற்கையாகப் பயன்படுத்தும் தமிழ், ஆங்கிலம் உள்ளிட்ட மொழிகளை கணினி புரிந்துகொண்டு, பகுத்தாய்ந்து, பொருள் உணர்ந்து, அதற்கேற்ப செயல்படச் செய்வதே இயற்கை மொழி செயலாக்கத்தின் அடிப்படை நோக்கமாகும். மனித–கணினி தொடர்பை எளிமைப்படுத்துவதிலும், மொழிபெயர்ப்பு, உரையிலிருந்து பேச்சு உருவாக்குதல், பேச்சிலிருந்து உரை உருவாக்குதல், உணர்வு பகுப்பாய்வு, கேள்வி–பதில் அமைப்புகள், தகவல் தேடல், உரைச் சுருக்கம் போன்ற பல்வேறு துறைகளிலும் NLP பயன்படுகிறது. தமிழ் போன்ற வளமான இலக்கிய மரபைக் கொண்ட மொழிகளுக்கு இயற்கை மொழி செயலாக்கம் புதிய வாய்ப்புகளை உருவாக்குகிறது. தமிழ் மொழியின் எழுத்து, சொல் அமைப்பு, இலக்கணம், விகுதி அமைப்பு, சொற்பொருள் வேறுபாடுகள், பேச்சுவழக்கு வேறுபாடுகள் போன்ற தனித்துவங்கள் NLP ஆய்வில் பல்வேறு சவால்களை உருவாக்குகின்றன. இக்கட்டுரை இயற்கை மொழி செயலாக்கத்தின் தோற்றம், வளர்ச்சி, அடிப்படைக் கூறுகள், பயன்பாடுகள், தமிழ்மொழியில் அதன் நிலை, சவால்கள் மற்றும் எதிர்கால வாய்ப்புகள் ஆகியவற்றை ஆராய்கிறது.

முக்கியச் சொற்கள்: இயற்கை மொழி செயலாக்கம், செயற்கை நுண்ணறிவு, கணினி மொழியியல், இயந்திர மொழிபெயர்ப்பு, தமிழ், உரை பகுப்பாய்வு, மொழித் தொழில்நுட்பம்.

முன்னுரை

மொழி என்பது மனிதனின் சிந்தனையையும் உணர்வுகளையும் அறிவையும் பிறரிடம் வெளிப்படுத்தும் முதன்மையான கருவியாகும். மனித சமூகத்தின் வளர்ச்சியோடு மொழியும் தொடர்ந்து மாற்றமடைந்து வந்துள்ளது. வாய்மொழி, எழுத்துமொழி, அச்சு ஊடகம், மின்னணு ஊடகம் என மொழியின் பயன்பாட்டுத் தளங்கள் விரிவடைந்துள்ளன. குறிப்பாகக் கணினி மற்றும் இணையத் தொழில்நுட்ப வளர்ச்சிக்குப் பின்னர் மொழி மற்றும் தொழில்நுட்பம் ஒன்றோடொன்று நெருக்கமாக இணைந்துள்ளன.

மனிதர்கள் கணினியுடன் தொடர்புகொள்ளும்போது, கணினி புரிந்துகொள்ளும் கட்டளைகளுக்கு மட்டுமே தங்களை வரையறுத்துக்கொள்ள வேண்டியிருந்தது. ஆனால் செயற்கை நுண்ணறிவின் வளர்ச்சியால் மனிதர்கள் பயன்படுத்தும் இயற்கை மொழியிலேயே கணினியுடன் தொடர்புகொள்ளும் வாய்ப்பு

உருவாகியுள்ளது. இந்த மாற்றத்தின் அடிப்படைத் தொழில்நுட்பமாக இயற்கை மொழி செயலாக்கம் விளங்குகிறது.“மனித மொழியை இயந்திரம் புரிந்துகொள்ளச் செய்வதே இயற்கை மொழி செயலாக்கத்தின் மைய நோக்கம்” எனக் கூறலாம்.

இயற்கை மொழி செயலாக்கம் என்பது மொழியியல், கணினி அறிவியல், செயற்கை நுண்ணறிவு, புள்ளியியல், தரவியல் போன்ற பல்வேறு அறிவுத்துறைகளின் சந்திப்புப் புள்ளியாகத் திகழ்கிறது. இன்று நாம் பயன்படுத்தும் தேடுபொறிகள், குரல் உதவியாளர்கள், தானியங்கி மொழிபெயர்ப்புகள், உரையாடல் செயற்கை நுண்ணறிவு அமைப்புகள் போன்றவற்றின் பின்னணியில் NLP தொழில்நுட்பம் செயல்படுகிறது.

இயற்கை மொழி செயலாக்கம் – விளக்கம்

Natural Language Processing (NLP) என்பதற்குத் தமிழில் “**இயற்கை மொழி செயலாக்கம்**” என்று பெயரிடப்படுகிறது. மனிதர்கள் இயல்பாகப் பயன்படுத்தும் மொழியை கணினி மூலம் செயலாக்குதல், பகுப்பாய்வு செய்தல், பொருள் கண்டறிதல் மற்றும் உருவாக்குதல் ஆகிய செயல்முறைகளின் தொகுப்பே NLP ஆகும். மனித மொழி கணினிக்கு நேரடியாகப் புரியாது. மனிதன் “நாளை சென்னைக்குச் செல்கிறேன்” என்று கூறும்போது, அந்த வாக்கியத்தில் உள்ள சொற்களின் பொருளையும், அவற்றுக்கிடையேயான தொடர்பையும், காலத்தையும், செயல் மற்றும் இடத்தையும் மனித மூளை இயல்பாகப் புரிந்துகொள்கிறது. ஆனால் கணினி அந்த வாக்கியத்தைத் தரவாகவே பார்க்கிறது. அந்தத் தரவுக்குப் பொருள் வழங்குவதற்காக மொழியியல் விதிகள், புள்ளியியல் மாதிரிகள், இயந்திரக் கற்றல், ஆழ்கற்றல் போன்ற தொழில்நுட்பங்கள் பயன்படுத்தப்படுகின்றன.

இயற்கை மொழி செயலாக்கத்தின் தோற்றமும் வளர்ச்சியும்

இயற்கை மொழி செயலாக்கத்தின் ஆரம்பகால வளர்ச்சி கணினி மொழியியல் மற்றும் இயந்திர மொழிபெயர்ப்பு ஆய்வுகளோடு தொடர்புடையது. கணினி வளர்ச்சியின் தொடக்கக் காலத்திலேயே மனித மொழியை இயந்திரம் புரிந்துகொள்ளச் செய்வது குறித்த ஆய்வுகள் தொடங்கின. ஆரம்பகால NLP அமைப்புகள் பெரும்பாலும் விதி அடிப்படையிலான அணுகுமுறையை (Rule-based approach) பின்பற்றின. மொழியின் இலக்கண விதிகள், சொற்களஞ்சியம், வாக்கிய அமைப்பு போன்றவற்றை மனிதர்கள் கணினியில் நிரலாக்கினர். பின்னர் அதிக அளவிலான மொழித் தரவுகள் கிடைக்கத் தொடங்கியதால் புள்ளியியல் அடிப்படையிலான மொழி செயலாக்கம் வளர்ச்சி பெற்றது. அதன் பின்னர் இயந்திரக் கற்றல் மற்றும் ஆழ்கற்றல் ஆகியவை NLP துறையில் பெரும் மாற்றத்தை ஏற்படுத்தின. இன்றைய காலகட்டத்தில் Transformer போன்ற ஆழ்கற்றல் கட்டமைப்புகள் மிகப்பெரிய அளவிலான மொழித் தரவுகளைக் கற்றுக்கொண்டு,

மனித மொழியைப் புரிந்துகொள்ளவும் உருவாக்கவும் உதவுகின்றன. இதன் தொடர்ச்சியாக பெரிய மொழி மாதிரிகள் உருவாகியுள்ளன. இதனால் NLP இன்று ஆய்வகங்களில் மட்டுமே காணப்படும் தொழில்நுட்பமாக இல்லாமல், கல்வி, மருத்துவம், வணிகம், நிர்வாகம், ஊடகம், வாடிக்கையாளர் சேவை உள்ளிட்ட பல துறைகளில் பயன்படுத்தப்படும் நடைமுறைத் தொழில்நுட்பமாக மாறியுள்ளது.

இயற்கை மொழி செயலாக்கத்தின் அடிப்படைக் கூறுகள்

இன்றைய அறிவியல் மற்றும் தொழில்நுட்ப வளர்ச்சியில் செயற்கை நுண்ணறிவு முக்கியப் பங்கு வகிக்கிறது. அதன் ஒரு முக்கியப் பிரிவாக இயற்கை மொழி செயலாக்கம் (Natural Language Processing – NLP) விளங்குகிறது. மனிதர்கள் பேசும் மற்றும் எழுதும் மொழிகளை கணினிகள் புரிந்துகொள்ளவும், அவற்றை ஆராயவும், பொருள் உணர்ந்து பதிலளிக்கவும் உதவும் தொழில்நுட்பமே இயற்கை மொழி செயலாக்கம் ஆகும்.

இயற்கை மொழி செயலாக்கத்தின் பொருள்

இயற்கை மொழி செயலாக்கம் என்பது தமிழ், ஆங்கிலம் போன்ற மனித மொழிகளை கணினிகள் புரிந்துகொண்டு செயல்பட உதவும் செயற்கை நுண்ணறிவுத் தொழில்நுட்பமாகும். மொழிபெயர்ப்பு, குரல் அடையாளம் காணுதல், உரையாடல் அமைப்புகள் மற்றும் உரை பகுப்பாய்வு போன்றவற்றில் இது பயன்படுகிறது.

இயற்கை மொழி செயலாக்கத்தின் அடிப்படைக் கூறுகள்

இயற்கை மொழி செயலாக்கம் என்பது மனிதர்கள் பயன்படுத்தும் இயற்கை மொழியை கணினிகள் புரிந்துகொள்ளவும், பகுப்பாய்வு செய்யவும், உருவாக்கவும் உதவும் செயற்கை நுண்ணறிவுத் தொழில்நுட்பமாகும். இதன் முக்கிய அடிப்படைக் கூறுகள் பின்வருமாறு,

ஒலியியல்

மொழியில் உள்ள ஒலிகள், அவற்றின் தன்மைகள் மற்றும் ஒலிகளுக்கிடையிலான தொடர்புகளைப் பற்றி ஆராய்வது ஒலியியல் ஆகும். பேச்சு வழியாக வழங்கப்படும் மொழியை கணினி புரிந்துகொள்ள இது உதவுகிறது.

உருவியல்

சொற்களின் அமைப்பு மற்றும் அவை உருவாகும் முறையை ஆராய்வது உருவியல் ஆகும். ஒரு சொல்லின் வேர் மற்றும் அதனுடன் சேரும் முன்னொட்டு, பின்னொட்டு போன்றவற்றை அடையாளம் காண்பது இதன் பணியாகும்.

தொடரியல்

சொற்கள் ஒன்றுடன் ஒன்று இணைந்து வாக்கியங்களை உருவாக்கும் முறையைத் தொடரியல் குறிக்கிறது. வாக்கியத்தில் சொற்களின் வரிசை மற்றும் இலக்கண அமைப்பைப் புரிந்துகொள்ள இது உதவுகிறது.

பொருளியல்

சொற்கள், சொற்றொடர்கள் மற்றும் வாக்கியங்களின் பொருளை ஆராய்வது பொருளியல் ஆகும். ஒரு வாக்கியம் உணர்த்தும் உண்மையான கருத்தைப் புரிந்துகொள்வதில் இது முக்கியப் பங்கு வகிக்கிறது.

பயன்பாட்டியல்

ஒரு மொழி குறிப்பிட்ட சூழலில் எவ்வாறு பயன்படுத்தப்படுகிறது என்பதைப் பயன்பாட்டியல் ஆராய்கிறது. பேசுபவரின் நோக்கம், சூழ்நிலை மற்றும் மறைமுகப் பொருள் ஆகியவற்றைப் புரிந்துகொள்ள இது உதவுகிறது.

சொற்பொருள் தெளிவாக்கம்

ஒரே சொல்லுக்குப் பல பொருள்கள் இருக்கும்போது, அது பயன்படுத்தப்பட்டுள்ள சூழலுக்கு ஏற்ற பொருளைக் கண்டறிவது இதன் பணியாகும்.

உரையாடல் பகுப்பாய்வு

பல வாக்கியங்களுக்கு இடையிலான தொடர்பைப் புரிந்துகொள்வது உரையாடல் பகுப்பாய்வு ஆகும். ஒரு உரையின் முழுமையான கருத்தையும் வாக்கியங்களுக்கிடையிலான தொடர்பையும் அறிய இது உதவுகிறது.

இயற்கை மொழி செயலாக்கத்தின் பயன்பாடுகள்

இயற்கை மொழி செயலாக்கம் பல்வேறு துறைகளில் பயன்படுத்தப்படுகிறது.

- இயந்திர மொழிபெயர்ப்பு: ஒரு மொழியிலிருந்து மற்றொரு மொழிக்கு மொழிபெயர்த்தல்.
- குரல் உதவியாளர்கள்: மனிதர்களின் குரல் கட்டளைகளைப் புரிந்துகொண்டு பதிலளித்தல்.
- உரை திருத்தம்: எழுத்துப் பிழைகள் மற்றும் இலக்கணப் பிழைகளைக் கண்டறிதல்.
- உணர்வு பகுப்பாய்வு: ஒரு கருத்து நேர்மறையானதா, எதிர்மறையானதா என்பதை அறிதல்.
- சாட்பாட்கள்: மனிதர்களுடன் உரையாடி அவர்களின் கேள்விகளுக்குப் பதிலளித்தல்.
- தகவல் தேடல்: ஆவணங்கள் மற்றும் இணையதளங்களில் தேவையான தகவல்களைக் கண்டறிதல்.

இயற்கை மொழி செயலாக்கத்தின் நன்மைகள்

இயற்கை மொழி செயலாக்கம் மனிதர்களுக்கும் கணினிகளுக்கும் இடையிலான தொடர்பை எளிதாக்குகிறது. மொழித் தடைகளைக் குறைக்கிறது. அதிக அளவிலான தகவல்களை விரைவாகப் பகுப்பாய்வு செய்ய உதவுகிறது. கல்வி, மருத்துவம், வணிகம் மற்றும் தகவல் தொழில்நுட்பம் போன்ற துறைகளில் நேரத்தையும் உழைப்பையும் மிச்சப்படுத்துகிறது.

தமிழ்மொழியில் NLP-யின் பயன்பாடுகள்

இயற்கை மொழி செயலாக்கம் என்பது மனிதர்கள் பயன்படுத்தும் இயற்கை மொழியை கணினி புரிந்துகொண்டு, பகுத்து, செயலாக்குவதற்கான செயற்கை நுண்ணறிவு தொழில்நுட்பமாகும். தமிழ்மொழியிலும் NLP முக்கியமான பங்காற்றி வருகிறது. மனிதர்கள் தமிழில் பேசுவதையும் எழுதுவதையும் கணினி புரிந்துகொள்ளவும், அதற்கேற்ற பதிலை வழங்கவும் NLP உதவுகிறது. தமிழிலிருந்து ஆங்கிலம் போன்ற பிற மொழிகளுக்கும், பிற மொழிகளிலிருந்து தமிழுக்கும் மொழிபெயர்ப்பு செய்வதில் NLP பயன்படுகிறது. மேலும், தமிழில் பேசப்படும் வார்த்தைகளை எழுத்து வடிவமாக மாற்றும் குரல் அறிதல் தொழில்நுட்பத்திலும் இது பயன்படுத்தப்படுகிறது. அதேபோல், தமிழில் உள்ள எழுத்து மற்றும் சொற்களை கணினி குரலாக மாற்றும் தொழில்நுட்பமும் NLP மூலம் செயல்படுகிறது. தமிழில் கேட்கப்படும் கேள்விகளுக்கு தானாக பதிலளிக்கும் Chatbot மற்றும் மெய்நிகர் உதவியாளர்களை உருவாக்குவதிலும் NLP பயன்படுகிறது. இணையதளங்கள் மற்றும் சமூக வலைத்தளங்களில் வெளியிடப்படும் தமிழ் கருத்துகளை ஆய்வு செய்து, அவை நேர்மறையானவையா அல்லது எதிர்மறையானவையா என்பதை கண்டறியவும் NLP உதவுகிறது. மேலும், நீண்ட தமிழ் கட்டுரைகள் மற்றும் ஆவணங்களில் உள்ள முக்கிய கருத்துகளை மட்டும் சுருக்கமாக வழங்கவும், தேவையான தகவல்களை விரைவாகத் தேடிக் கண்டுபிடிக்கவும் இது பயன்படுகிறது. கல்வித் துறையில் தமிழ் இலக்கணப் பிழைகளை கண்டறிதல், மொழி கற்றல் மற்றும் தானியங்கி மதிப்பீடு போன்றவற்றிலும் NLP பயன்படுத்தப்படுகிறது. இதன் மூலம் தமிழ்மொழியை நவீன கணினி மற்றும் செயற்கை நுண்ணறிவு தொழில்நுட்பங்களுடன் இணைக்க முடிகிறது. எனவே, கல்வி, தகவல் தொடர்பு, வணிகம், ஊடகம் போன்ற பல்வேறு துறைகளில் தமிழ்மொழியின் பயன்பாட்டை மேம்படுத்துவதில் NLP முக்கியத்துவம் வாய்ந்ததாக உள்ளது.

முடிவுரை

இயற்கை மொழி செயலாக்கம் என்பது கணினி அறிவியலின் ஒரு தொழில்நுட்பப் பிரிவு மட்டுமல்ல; அது மனித மொழிக்கும் இயந்திர அறிவுக்கும் இடையே பாலமாகச் செயல்படும் ஒரு முக்கிய அறிவியல் துறையாகும். மனிதன் சிந்திப்பதைப் போல

கணினி சிந்திக்கிறது என்று கூறுவதற்குப் பதிலாக, மனிதன் பயன்படுத்தும் மொழியை கணினி புரிந்துகொண்டு அதற்கேற்ற வகையில் செயல்படச் செய்வதே NLP-யின் முக்கிய நோக்கமாகும். மொழிபெயர்ப்பு முதல் குரல் தொழில்நுட்பம் வரை, தேடுபொறி முதல் உரையாடல் செயற்கை நுண்ணறிவு வரை, கல்வி முதல் இலக்கிய ஆய்வு வரை NLP-யின் பயன்பாட்டுப் பரப்பு விரிவடைந்து கொண்டிருக்கிறது.

தமிழ்மொழியைப் பொறுத்தவரை, அதன் தொன்மை மற்றும் இலக்கிய வளத்தை நவீன தொழில்நுட்பத்துடன் இணைக்கும் மிக முக்கியமான வாய்ப்பாக NLP அமைகிறது. தமிழ் நூல்களை மின்னாக்கம் செய்தல், மொழித் தரவுத்தொகுப்புகளை உருவாக்குதல், தமிழ் இலக்கணத்தை அடிப்படையாகக் கொண்ட மொழி மாதிரிகளை உருவாக்குதல், தமிழ் பேச்சுத் தரவுகளைப் பாதுகாத்தல் ஆகிய பணிகள் எதிர்காலத்தில் முக்கியத்துவம் பெறுகின்றன. எனவே, “மொழியின் வளத்தைத் தொழில்நுட்பத்தின் வலிமையுடன் இணைக்கும் புதிய அறிவியல் பாலமே இயற்கை மொழி செயலாக்கம் எனக் குறிப்பிடலாம்.

துணைநூற்பட்டியல்

References

அகத்தியலிங்கம், ச. — *மொழியியல்*, அண்ணாமலைப் பல்கலைக்கழகம், சிதம்பரம்.

சண்முகம், செ. வை. — *தமிழ் மொழி அமைப்பும் மொழியியலும்*, மணிவாசகர் பதிப்பகம், சென்னை.

பொற்கோ — *தமிழ் மொழியியல்*, உலகத் தமிழாராய்ச்சி நிறுவனம், சென்னை.

இராமசாமி, கி. — *கணினித் தமிழ்*, தமிழ் இணையக் கல்விக்கழகம், சென்னை.

தமிழண்ணல் — *தமிழ் இலக்கிய வரலாறு*, மீனாட்சி புத்தக நிலையம், மதுரை.