



தமிழ்மணம் சர்வதேசத் தமிழ் ஆய்விதழ்

Peer-Reviewed | Open Access | Crossref | Google Scholar
SJIF Impact Factor 3.5 | Multidisciplinary

Article DOI: <https://doi.org/10.63300/tm12022026.43>



Tamil Literary Research through Artificial Intelligence (AI) and Prompt Engineering: New Research Approaches

Dr. R.Kannan^{1*}, Dr M.Ramachandran², Dr.V.C.Srinivasan³,

¹Assistant Professor, Department of Tamil and other Languages,
RVS College of Arts & Science, Sulur, Coimbatore - 641 402.

²Associate Professor In Tamil, Nandha Arts And Science College (Autonomous), Erode.

Email: rampiththan@gmail.com

Orcid id: <https://orcid.org/0009-0000-3815-4580>

³Administrative Officer, Nandha Arts and Science College (Autonomous), Erode.

Email: srvc2345@gmail.com, Orcid Id: <https://orcid.org/0009-0008-9926-7982>.

*Corresponding Author Email: vaanavilkannan@gmail.com.

Article Info

Received on 27-June-2026, Revised on 29-June-2026, Accepted on 29-June-2026, Published on 20-July-2026

ABSTRACT

Tamil literary research has traditionally been rooted in textual analysis, comparative study, and historical approaches. However, recent advancements in Large Language Models (LLMs) and Prompt Engineering techniques have introduced new tools and methodologies to the field of Tamil literary studies. This paper explores how artificial intelligence and Prompt Engineering techniques can be applied to Tamil literary research. Specifically, focusing on language models, datasets, and evaluation metrics developed for Tamil within the context of low-resource languages, this paper describes how these can be utilized for tasks such as text analysis, sentiment analysis, summarization, translation, and question-answering systems. Additionally, the paper highlights the challenges, ethical concerns, and the necessity of human expertise when adopting these technologies.

KEYWORDS: Artificial Intelligence, Prompt Engineering, Tamil Literature, Large Language Model, Computational Tamilology.



Copyright © 2024 by the author(s). Published by Department of Library, Nallamuthu Gounder Mahalingam College, Pollachi. This is an open access article under the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).



செயற்கை நுண்ணறிவு (AI) மற்றும் Prompt Engineering வழி தமிழ் இலக்கிய ஆய்வு: புதிய ஆராய்ச்சி அணுகுமுறைகள்

¹முனைவர் அர.கண்ணன், உதவிப் பேராசிரியர்., தமிழ் மற்றும் பிற மொழிகள் துறை,

ஆர்.வி.எஸ் கலை மற்றும் அறிவியல் கல்லூரி, சூலூர், கோயம்புத்தூர்.

²முனைவர் மா. இராமச்சந்திரன், இணைப் பேராசிரியர், தமிழ்த் துறை, நந்தா கலை மற்றும்
அறிவியல் கல்லூரி (தன்னாட்சி)-ஈரோடு.

³முனைவர் வி.சி.சீனிவாசன், நிர்வாக அதிகாரி,

நந்தா கலை மற்றும் அறிவியல் கல்லூரி (தன்னாட்சி), ஈரோடு.

ஆய்வுச் சுருக்கம்

தமிழ் இலக்கிய ஆய்வு பாரம்பரியமாக உரையாய்வு, ஒப்பீட்டு ஆய்வு, வரலாற்று அணுகுமுறை ஆகியவற்றை அடிப்படையாகக் கொண்டு இயங்கி வந்துள்ளது. எனினும், சமீப காலமாக பெருமொழி மாதிரிகள் (Large Language Models) மற்றும் Prompt Engineering நுட்பங்களின் வளர்ச்சி, தமிழ் இலக்கிய ஆய்வுக்கும் புதிய கருவிகளையும் அணுகுமுறைகளையும் வழங்கியுள்ளது. இக்கட்டுரை, செயற்கை நுண்ணறிவும் Prompt Engineering நுட்பங்களும் தமிழ் இலக்கிய ஆய்வில் எவ்வாறு பயன்படுத்தப்படலாம் என்பதை ஆராய்கிறது. குறிப்பாக, குறைவளமொழி (Low-Resource Language) நிலையில் உள்ள தமிழுக்கென உருவாக்கப்பட்ட மொழி மாதிரிகள், தரவுத்தொகுப்புகள், மற்றும் மதிப்பீட்டு அளவுகோல்கள் ஆகியவற்றை மையமாகக் கொண்டு, உரைப் பகுப்பாய்வு, உணர்ச்சி அடையாளங்காணல், சுருக்கமாக்கல், மொழிபெயர்ப்பு, கேள்வி-பதில் அமைப்புகள் போன்ற பணிகளுக்கு இவை எவ்வாறு பயன்படுத்தப்படலாம் என்பதை இக்கட்டுரை விவரிக்கிறது. மேலும், இத்தகைய தொழில்நுட்பங்களைப் பயன்படுத்தும்போது எழும் சவால்களையும், நெறிமுறைக் கவலைகளையும், மனித நிபுணத்துவத்தின் தேவையையும் இக்கட்டுரை எடுத்துரைக்கிறது.

திறவுச் சொற்கள்: செயற்கை நுண்ணறிவு, Prompt Engineering, தமிழ் இலக்கியம், பெருமொழி மாதிரி, கணினித் தமிழியல்.

1. முன்னுரை

தமிழ் இலக்கியம் இரண்டாயிரம் ஆண்டுகளுக்கும் மேலான தொடர்ச்சியான இலக்கிய மரபைக் கொண்டது; சங்க இலக்கியம் முதல் நவீன இலக்கியம் வரையிலான இப்பரந்த தொகுப்பு, மொழியியல், சமூகவியல், வரலாறு, தத்துவம் ஆகிய பல்வேறு தளங்களில் ஆய்வுக்கு உட்படுத்தப்பட்டு வருகிறது. பாரம்பரிய இலக்கிய ஆய்வு முறைகள், கையேடு உரையாய்வு, கையால் குறிப்பெடுத்தல், நிபுணர் விளக்கவுரை ஆகியவற்றை அடிப்படையாகக் கொண்டவை;

இவை ஆழமான புரிதலை வழங்கினாலும், பெருமளவிலான உரைத் தொகுப்புகளை விரைவாக பகுப்பாய்வு செய்வதில் வரம்புகளைக் கொண்டுள்ளன.

இக்குறையை நிவர்த்தி செய்யும் வகையில், சமீப காலமாக செயற்கை நுண்ணறிவு தொழில்நுட்பம், குறிப்பாக பெருமொழி மாதிரிகள் (Large Language Models – LLMs), இயற்கை மொழி ஆய்வுத் துறையில் (Natural Language Processing) பெரும் மாற்றத்தை ஏற்படுத்தியுள்ளது (Sahoo et al., 2024). இம்மாதிரிகளை திறம்பட பயன்படுத்துவதற்கான நுட்பமாக Prompt Engineering வளர்ச்சியடைந்துள்ளது இது மாதிரியின் அடிப்படை அளவுருக்களை மாற்றாமலேயே, சரியான வழிமுறை வாக்கியங்களை (Prompts) வடிவமைப்பதன் மூலம் விரும்பிய வெளியீட்டைப் பெறும் நுட்பமாகும் (Vatsal & Dubey, 2024). தமிழைப் பொறுத்தவரை, குறைவளமொழி என்ற நிலையில் இருந்தாலும், தமிழ்-குறிப்பிட்ட மொழி மாதிரிகள் (Balachandran, 2023) மற்றும் இந்திய மொழிகளுக்கான பரந்த கணினித் தமிழியல் தரவுத்தொகுப்புகள் (Doddapaneni et al., 2023) சமீப காலமாக உருவாக்கப்பட்டு வருகின்றன.

இத்தகைய தொழில்நுட்ப வளர்ச்சிகள், தமிழ் இலக்கிய ஆய்வாளர்களுக்கு புதிய வாய்ப்புகளைத் திறந்துவிடுகின்றன பாரம்பரிய உரையாய்வு முறைகளுக்கு மாற்றாக அல்ல, மாறாக அவற்றை நிறைவு செய்யும் கருவிகளாக. இக்கட்டுரை, இத்தகைய கருவிகள் தமிழ் இலக்கிய ஆய்வில் எவ்வாறு பயன்படுத்தப்படலாம், அவற்றின் சாத்தியக்கூறுகள் என்ன, மற்றும் அவற்றைப் பயன்படுத்தும்போது கருத்தில் கொள்ளவேண்டிய வரம்புகள் யாவை என்பதை விரிவாக ஆராய்கிறது.

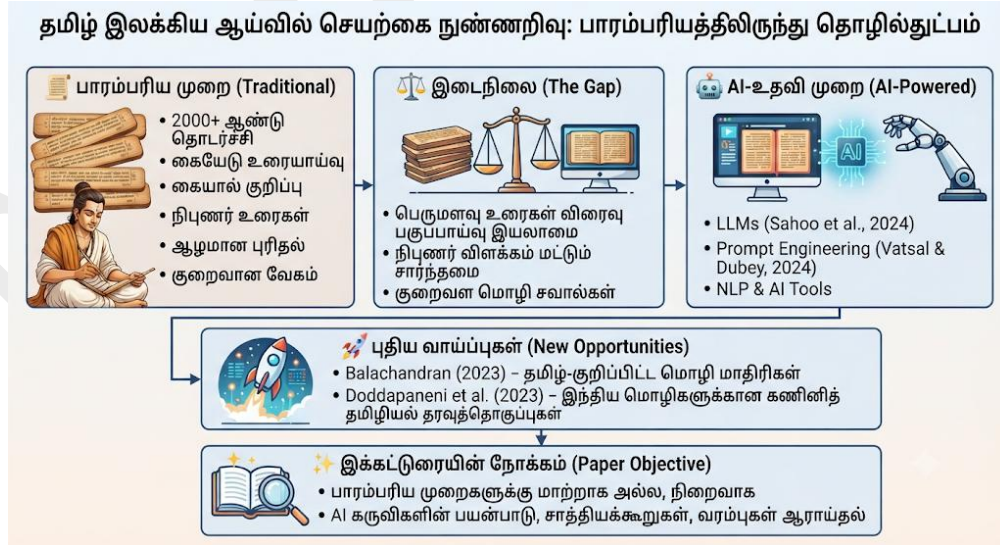


Fig1: தமிழ் இலக்கிய ஆய்வில் செயற்கை நுண்ணறிவு: பாரம்பரியத்திலிருந்து தொழில்நுட்பம்

இப்படம் தமிழ் இலக்கிய ஆய்வில் செயற்கை நுண்ணறிவின் (AI) பங்கை விளக்குகிறது. இது 2000 ஆண்டுகளுக்கும் மேலான பாரம்பரிய முறைகளையும், தற்காலச் சவால்களையும் தாண்டி, பெரிய மொழி மாதிரிகள் (LLMs) மற்றும் நவீன கருவிகள் மூலம் ஆய்வுகளை எவ்வாறு மேம்படுத்தலாம் என்பதைத் தெளிவாகக் காட்டுகிறது. பாரம்பரியத்திற்கு மாற்றாக அல்லாமல், அவற்றை நிறைவு செய்யும் வகையில் AI உதவுகிறது.

2. ஆய்வின் நோக்கங்கள்

1. தமிழ் மொழிக்கான செயற்கை நுண்ணறிவு தொழில்நுட்ப வளர்ச்சியை மொழி மாதிரிகள், தரவுத்தொகுப்புகள், மதிப்பீட்டு அளவுகோல்கள் என்ற வகையில் ஆவணப்படுத்துதல்.

Prompt Engineering எனும் நுட்பத்தின் அடிப்படைக் கருத்தாக்கத்தையும் அதன் முக்கிய வகைகளையும் விளக்குதல்.

3. தமிழ் இலக்கிய ஆய்வில் AI மற்றும் Prompt Engineering பயன்படக்கூடிய குறிப்பிட்ட பணிகளை உரைப் பகுப்பாய்வு, உணர்ச்சி அடையாளங்காணல், சுருக்கமாக்கல், மொழிபெயர்ப்பு, கேள்வி-பதில் அமைப்பு என்ற வகையில் அடையாளம் காணுதல்.

இத்தொழில்நுட்பங்களைப் பயன்படுத்துவதில் உள்ள சவால்களையும் நெறிமுறைக் கவலைகளையும் மதிப்பிடுதல்.

தமிழ் இலக்கிய ஆய்வுக்கும் கணினித் தமிழியலுக்கும் இடையேயான எதிர்கால ஒத்துழைப்புக்கான திசைகளை முன்மொழிதல்.

3. ஆய்வு முறையியல்

இக்கட்டுரை தரமுறை ஆவணப் பகுப்பாய்வு (Qualitative Documentary Analysis) அணுகுமுறையைப் பின்பற்றுகிறது. தமிழ் மொழிக்கான கணினித் தமிழியல் தொடர்பான சமீபத்திய ஆய்வுக் கட்டுரைகள் (Balachandran, 2023; Chakravarthi et al., 2022; Doddapaneni et al., 2023; Varsha et al., 2025), பொதுவான Prompt Engineering நுட்பங்கள் தொடர்பான ஆய்வுகள் (Sahoo et al., 2024; Vatsal & Dubey, 2024; Minaee et al., 2024), மற்றும் தமிழ் இலக்கிய ஆய்வில் AI பயன்பாடு தொடர்பான ஆய்வுக் கட்டுரைகள் (Veerakannan & Vijayakumar, 2025; Veerakannan, 2025) ஆகியவை முதன்மைத் தரவாக ஆய்வு செய்யப்பட்டுள்ளன. மேலும், தமிழ் இலக்கியத்தை பன்முகத் துறை அணுகுமுறையில் (Interdisciplinary Approach) ஆய்வு செய்யும் சமீபத்திய முன்னுதாரண ஆய்வுகளும் (Ramesh & Thamilarasi, 2025; Brindha, 2024; Sopa & Murugeswari, 2023; Raja, 2022) ஒப்பீட்டு நோக்கில் இணைக்கப்பட்டுள்ளன, ஏனெனில் இவை தமிழ் இலக்கியத்தை நவீன அறிவியல் கண்ணோட்டங்களுடன் இணைத்து ஆய்வு செய்யும்

பொதுவான போக்கை எடுத்துக்காட்டுகின்றன.

4. தமிழ் இலக்கிய ஆய்வின் பாரம்பரிய அணுகுமுறைகள்

பாரம்பரிய தமிழ் இலக்கிய ஆய்வு, மூலநூல் உரையாய்வு, உரையாசிரியர் விளக்கங்களை ஒப்பிடுதல், சொல்லாராய்ச்சி, யாப்பு-அணி பகுப்பாய்வு ஆகியவற்றை அடிப்படையாகக் கொண்டது. இம்முறைகள், நிபுணத்துவம் மிக்க ஆய்வாளர்களால் பல தசாப்தங்களாக மேற்கொள்ளப்பட்டு, தமிழ் இலக்கிய ஆய்வுக்கு உறுதியான அடித்தளத்தை அமைத்துள்ளன. எனினும், சங்க இலக்கியம், பதினெண்கீழ்க்கணக்கு, பக்தி இலக்கியம் என பரந்துபட்ட நூற்றுக்கணக்கான நூல்களை, சொல் அளவிலான அமைப்பு முறைகள், தொடர்ச்சியான கருப்பொருள் போக்குகள், அல்லது எழுத்தாளர் பாணி ஒப்பீடு ஆகியவற்றின் அடிப்படையில் பகுப்பாய்வு செய்வதென்பது, கையால் மட்டுமே செய்யும்போது கால அளவிலும் நோக்கு அளவிலும் வரம்புகளைக் கொண்டிருக்கிறது. இக்குறையை நிவர்த்தி செய்யும் இடத்திலேயே கணினி வழி உரைப் பகுப்பாய்வு நுட்பங்களின் பங்களிப்பு முக்கியத்துவம் பெறுகிறது.

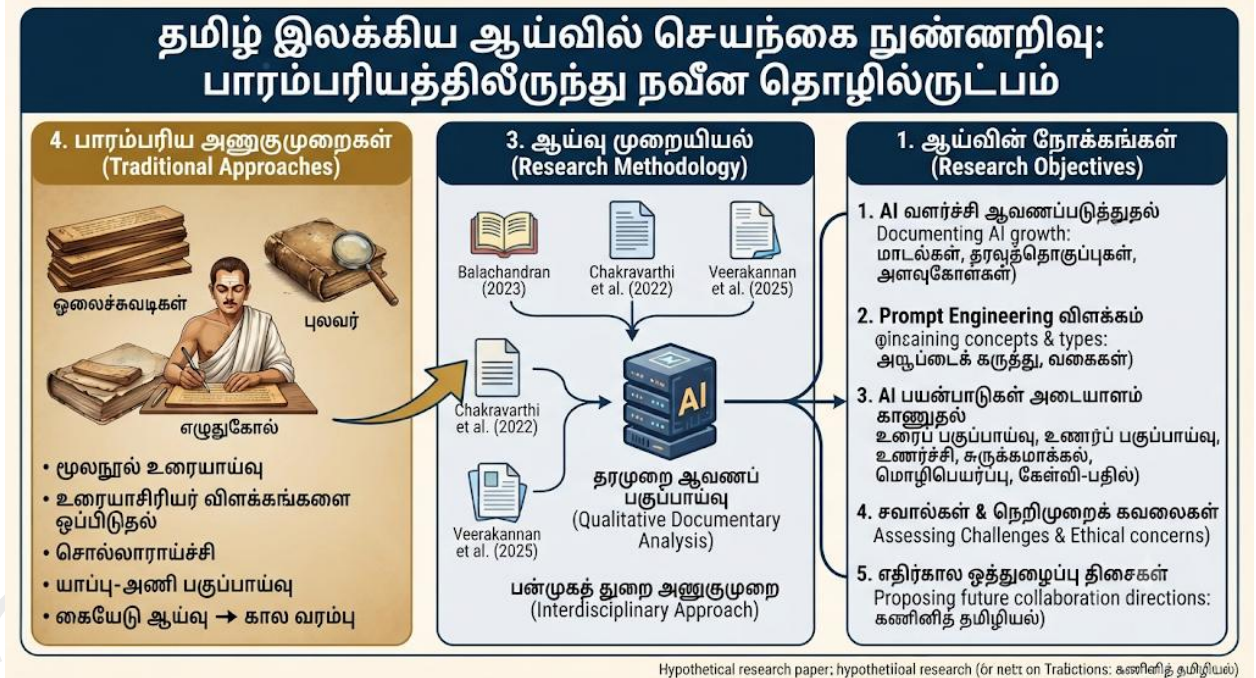


Fig2: தமிழ் இலக்கிய ஆய்வில் செயற்கை நுண்ணறிவு: பாரம்பரியத்திலிருந்து நவீன தொழில்நுட்பம்

இப்படப்படம் தமிழ் இலக்கிய ஆய்வில் பாரம்பரிய முறைகள் மற்றும் நவீன தொழில்நுட்பத்தின் ஒருங்கிணைப்பை விளக்குகிறது. ஓலைச்சுவடிகளை வைத்துப் புலவர்கள் ஆய்வு செய்யும் காலத்திலிருந்து, செயற்கை நுண்ணறிவைப் பயன்படுத்தி உரைப் பகுப்பாய்வு, மொழிபெயர்ப்பு மற்றும் தரவுத் தொகுப்பு உருவாக்கம் போன்ற நவீன ஆய்வுகளை நோக்கி நகர்வதை இது

காட்டுகிறது. இதன் மூலம், பன்முகத் துறை அணுகுமுறையைக் கையாண்டு எதிர்கால ஆய்வுகளுக்கான வழிகள் முன்மொழியப்படுகின்றன.

5. செயற்கை நுண்ணறிவும் தமிழ் மொழி தொழில்நுட்பமும்

தமிழ், மொழியியல் ரீதியாக வளமான ஒட்டு மொழியாக (Agglutinative Language) இருந்தபோதிலும், பெருமொழி மாதிரிகளின் பயிற்சித் தரவுத் தொகுப்புகளில் ஒப்பீட்டளவில் குறைவான பிரதிநிதித்துவத்தையே பெற்றுள்ளது; இதனால் தமிழ் ஒரு “குறைவளமொழி” (Low-Resource Language) எனக் கருதப்படுகிறது (Varsha et al., 2025). இக்குறையை நிவர்த்தி செய்யும் முயற்சியாக, பாலச்சந்திரன் (Balachandran, 2023) Llama 2 மாதிரியில் பதினாறாயிரம் தமிழ் டோக்கன்களைச் சேர்த்து Tamil-LLaMA எனும் தமிழுக்கெனத் தனிப்பயனாக்கப்பட்ட மொழி மாதிரியை உருவாக்கியுள்ளார்; இது தமிழ் உரை உருவாக்கத்திலும் புரிதலிலும் மேம்பட்ட திறனை வழங்குகிறது. இதேபோன்று, தொடப்பனேனி மற்றும் பலர் (Doddapaneni et al., 2023) இந்திய மொழிகளுக்கான ஒற்றைமொழி தரவுத் தொகுப்புகள், அளவுகோள்கள், மாதிரிகளை உருவாக்கி, தமிழ் உள்ளிட்ட இந்திய மொழிகள் பெருமொழி மாதிரி ஆய்வில் பின் தங்கிவிடாமல் இருக்க பங்களித்துள்ளனர்.

மேலும், சக்கரவர்த்தி மற்றும் பலர் (Chakravarthi et al., 2022) உருவாக்கிய DravidianCodeMix தரவுத் தொகுப்பு, தமிழ்-ஆங்கிலம் கலப்பு மொழி (Code-Mixed) உரைகளில் உணர்ச்சிப் பகுப்பாய்வு மற்றும் தகாத மொழி அடையாளங்காணல் ஆகியவற்றுக்கான முதன்மையான வளமாக விளங்குகிறது. இத்தகைய தரவுத் தொகுப்புகள், சமூக ஊடக தமிழ் உரைகளை பகுப்பாய்வு செய்வதற்கு மட்டுமல்லாமல், நவீன தமிழ் இலக்கிய அமைப்புகளில் (குறிப்பாக ஊடகவியல் சார்ந்த படைப்புகளில்) காணப்படும் கலப்பு மொழிப் பயன்பாட்டை ஆய்வு செய்வதற்கும் பயன்படக்கூடும். இருப்பினும், “இலக்கணம் முதல் பொருள் வரை” (Varsha et al., 2025) எனும் தலைப்பில் மேற்கொள்ளப்பட்ட சமீபத்திய ஆய்வு, மூடிய-மூல (Closed-Source) மற்றும் திறந்த-மூல (Open-Source) பெருமொழி மாதிரிகள் இரண்டும் தமிழ் இலக்கண நுணுக்கங்களைப் புரிந்துகொள்வதில் இன்னும் குறிப்பிடத்தக்க வரம்புகளைக் கொண்டுள்ளன என்பதை எடுத்துக்காட்டுகிறது இது தமிழ் இலக்கிய ஆய்வுக்கு இத்தொழில்நுட்பங்களைப் பயன்படுத்தும்போது கவனமாக இருக்க வேண்டிய அவசியத்தை உணர்த்துகிறது.

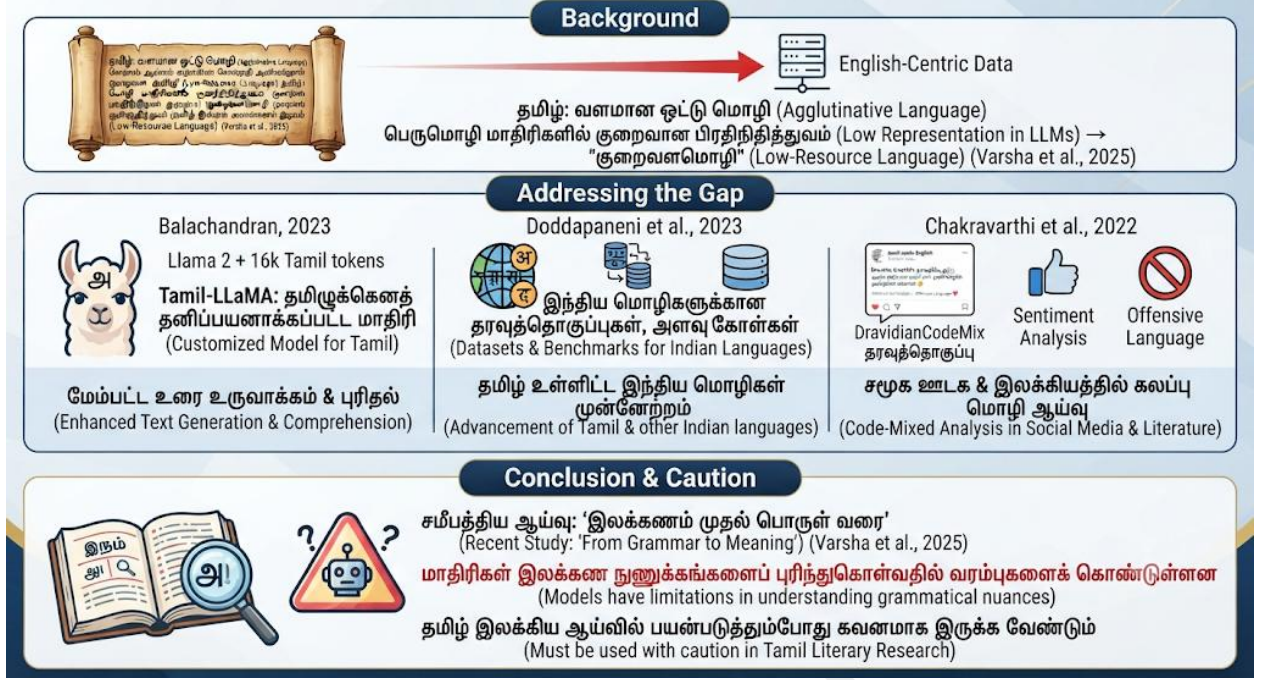


Fig3: பெருமொழி மாதிரிகளில் தமிழ்: சவால்களும் தீர்வுகளும்

தமிழ் மொழியியல் ரீதியாக வளம் மிக்க ஒட்டு மொழியாக இருந்தாலும், பெருமொழி மாதிரிகளின் பயிற்சித் தரவுகளில் குறைவாகவே உள்ளது. இதனால் இது ஒரு 'குறைவளமொழி' எனக் கருதப்படுகிறது. பாலச்சந்திரன் மற்றும் தொடப்பனேனி போன்றோரின் ஆய்வுகள், தனிப்பயனாக்கப்பட்ட மாதிரிகள் மற்றும் தரவுத் தொகுப்புகளை உருவாக்குவதன் மூலம் இக்குறையை நிவர்த்தி செய்கின்றன. இருப்பினும், சக்கரவர்த்தி குழுவின் ஆய்வு, நவீன சமூக ஊடக மற்றும் இலக்கியத் தரவுகளில் உள்ள கலப்பு மொழி பயன்பாட்டைப் பகுப்பாய்வு செய்வதற்கான வழிகளைக் காட்டுகிறது. சமீபத்திய ஆய்வு, மாதிரிகள் தமிழ் இலக்கண நுணுக்கங்களைப் புரிந்துகொள்வதில் இன்னும் குறிப்பிடத்தக்க வரம்புகளைக் கொண்டுள்ளன என்பதை வலியுறுத்துகிறது. எனவே, தமிழ் இலக்கிய ஆய்வுக்கு இவற்றைப் பயன்படுத்தும்போது எச்சரிக்கையுடன் இருக்க வேண்டும்.

6. Prompt Engineering: கருத்தாக்கமும் நுட்பங்களும்

Prompt Engineering என்பது, பெருமொழி மாதிரியின் அடிப்படை அளவுருக்களை மறுபயிற்சி செய்யாமலேயே, சரியாக வடிவமைக்கப்பட்ட வழிமுறை வாக்கியங்களின் (Prompts) மூலம் மாதிரியிடமிருந்து துல்லியமான, சூழல் பொருத்தமான வெளியீட்டைப் பெறும் நுட்பமாகும் (Sahoo et al., 2024). இதன் முக்கிய வகைகளாக பின்வருவனவற்றைக் குறிப்பிடலாம்:

ஜீரோ-ஷாட் ப்ரம்ப்டிங் (Zero-Shot Prompting) முன்னுதாரண எடுத்துக்காட்டுகள் எதுவும் இன்றி, நேரடியாக பணியை வழிமுறையாகக் கொடுத்தல்.

ஃபியூ-ஷாட் ப்ராம்ப்டிங் (Few-Shot Prompting) சில எடுத்துக்காட்டு உள்ளீடு-வெளியீடு ஜோடிகளை வழிமுறையுடன் இணைத்து அளித்தல், இதன் மூலம் மாதிரி எதிர்பார்க்கப்படும் அமைப்பையும் பாணியையும் புரிந்துகொள்கிறது.

சங்கிலி-சிந்தனை ப்ராம்ப்டிங் (Chain-of-Thought Prompting) மாதிரியை படிப்படியான தர்க்க ரீதியிலான பகுப்பாய்வு செய்யத் தூண்டுதல்; இலக்கிய உரை பகுப்பாய்வில் இது குறிப்பாக பயனுள்ளதாக இருக்கும், ஏனெனில் ஒரு பாடலின் பொருள் விளக்கத்தை படிநிலையாக நியாயப்படுத்த உதவுகிறது (Vatsal & Dubey, 2024).

பங்கு-அடிப்படையிலான ப்ராம்ப்டிங் (Role-Based Prompting) மாதிரியை ஒரு குறிப்பிட்ட நிபுணத்துவப் பாத்திரத்தில் (எ.கா., “தமிழ் இலக்கிய ஆய்வாளர்”) செயல்படும்படி வழிநடத்துதல், இதன் மூலம் பதில்களின் சூழல் பொருத்தமும் தரமும் மேம்படுகிறது.

இத்தகைய நுட்பங்கள், பொதுவான NLP பணிகளுக்கு (சுருக்கமாக்கல், மொழிபெயர்ப்பு, வகைப்படுத்தல் போன்றவை) பரவலாகப் பயன்படுத்தப்பட்டு வந்தாலும், தமிழ் இலக்கியம் போன்ற கலாச்சார ரீதியாக ஆழமான, செறிவான குறியீட்டு மொழியைக் கொண்ட உரைத் தொகுப்புகளுக்கு இவற்றைப் பொருத்தமாக வடிவமைப்பது ஒரு தனித்துவமான சவாலாகும் (Minaee et al., 2024).

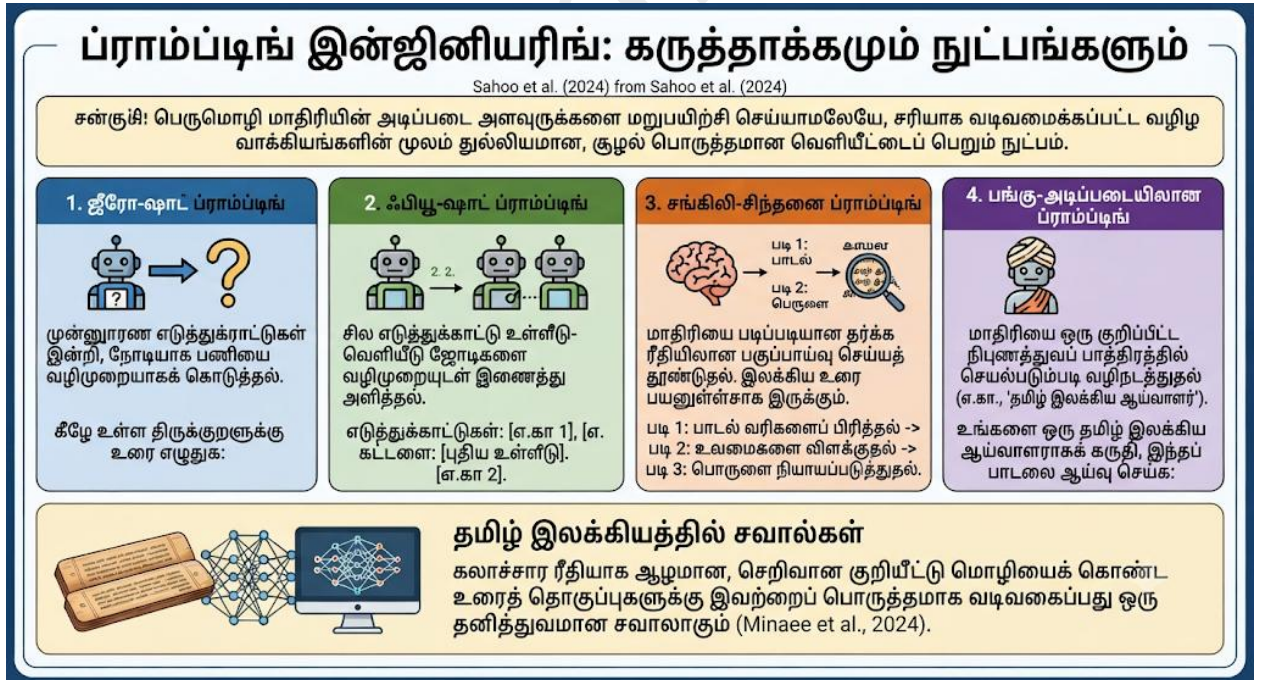


Fig4: ப்ராம்ப்டிங் இன்ஜினியரிங்: கருத்தாக்கமும் நுட்பங்களும்

இப்படம் பெருமொழி மாதிரிகளில் பயன்படுத்தப்படும் ப்ராம்ப்டிங் இன்ஜினியரிங் நுட்பங்களை விளக்குகிறது. ஜீரோ-ஷாட், பியூ-ஷாட், சங்கிலி-சிந்தனை மற்றும் பங்கு-அடிப்படையிலான

ப்ராம்படிங் முறைகள் மூலம் துல்லியமான வெளியீட்டைப் பெறுவதை இது காட்டுகிறது. இருப்பினும், தமிழ் இலக்கியம் போன்ற கலாச்சார ஆழம் கொண்ட உரைத்தொகுப்புகளுக்கு இவற்றை வடிவமைப்பது ஒரு தனித்துவமான சவாலாகும் என்பதை இது சுட்டிக்காட்டுகிறது.

7. தமிழ் இலக்கிய ஆய்வில் AI மற்றும் Prompt Engineering பயன்பாட்டு சாத்தியக்கூறுகள்

7.1 உரைப் பகுப்பாய்வும் கருப்பொருள் அடையாளங்காணலும்

பெருமொழி மாதிரிகளுக்கு பொருத்தமான Few-Shot Prompts வழங்குவதன் மூலம், சங்க இலக்கியம், பதினெண்கீழ்க்கணக்கு நூல்கள் போன்ற பெருந்தொகுப்புகளில் இடம்பெறும் கருப்பொருள்களை (எ.கா., அறம், அரசியல் நீதி, சமூக விழுமியங்கள்) தானியங்கு முறையில் வகைப்படுத்தலாம். இத்தகைய பகுப்பாய்வு அணுகுமுறை, ரமேஷ் மற்றும் தமிழரசி (Ramesh & Thamilarasi, 2025) புறநானூற்றை நடத்தை உளவியல் கோணத்தில் ஆய்வு செய்த முறையையும், பிருந்தா (Brindha, 2024) நற்றிணையில் மெய்ப்பாடுகளை பகுப்பாய்வு செய்த முறையையும் ஒத்திருக்கிறது ஆனால் அளவில் மிகப் பெரிய தொகுப்புகளுக்கு இதைப் பரவலாக்கும் திறனை AI வழங்குகிறது.

7.2 உணர்ச்சி அடையாளங்காணலும் பாணி பகுப்பாய்வும்

DravidianCodeMix போன்ற தரவுத்தொகுப்புகள் மூலம் பயிற்சி பெற்ற மாதிரிகள், தமிழ் இலக்கியப் பாடல்களில் இடம்பெறும் உணர்ச்சி நிலைகளை (எ.கா., அகப்பொருள் திணைகளில் இடம்பெறும் புணர்தல், பிரிதல், இருத்தல் ஆகிய உணர்வு நிலைகள்) அளவிடும் ஆய்வுகளுக்கு அடிப்படையாக அமையலாம் (Chakravarthi et al., 2022). மேலும், எழுத்தாளர் பாணி ஒப்பீடு (Stylometric Analysis) மூலம், ஒரு பனுவல் பல ஆசிரியர்களால் எழுதப்பட்டதா அல்லது ஒரு ஆசிரியரின் படைப்பே தொடர்ச்சியாக உள்ளதா என்பதை கணக்கிடும் ஆய்வுகளும் சாத்தியமாகின்றன.

7.3 சுருக்கமாகவும் மொழிபெயர்ப்பும்

பெருமொழி மாதிரிகள், நீண்ட உரையாசிரியர் விளக்கவுரைகளையும் பழந்தமிழ் நூல்களையும் நவீன தமிழிலும் ஆங்கிலத்திலும் சுருக்கமாக மொழிபெயர்க்க உதவக்கூடும் இது சர்வதேச ஆய்வாளர்களுக்கு தமிழ் இலக்கியத்தை அணுகும் வாய்ப்பை விரிவுபடுத்தும். எனினும், இத்தகைய மொழிபெயர்ப்புகள் இலக்கிய நுணுக்கங்களையும் கலாச்சாரக் குறியீடுகளையும் இழக்கும் அபாயம் உள்ளது என்பதால், மனித நிபுணர் மேற்பார்வை இன்றியமையாததாகும்.

7.4 கேள்வி-பதில் அமைப்புகளும் கல்விசார் பயன்பாடும்

Chain-of-Thought Prompting நுட்பத்தைப் பயன்படுத்தி, மாணவர்கள் மற்றும்

ஆய்வாளர்களுக்கு திருக்குறள், நாலடியார் போன்ற நூல்களின் குறிப்பிட்ட குறள்கள்/பாடல்கள் தொடர்பான கேள்விகளுக்கு படிநிலையான, நியாயப்படுத்தப்பட்ட விளக்கங்களை வழங்கும் கல்விசார் கருவிகளை உருவாக்க முடியும் (Veerakannan & Vijayakumar, 2025). இத்தகைய கருவிகள், தமிழ் இலக்கியக் கல்வியை மேலும் அணுகக்கூடியதாக மாற்றும் திறனைக் கொண்டுள்ளன (Veerakannan, 2025).

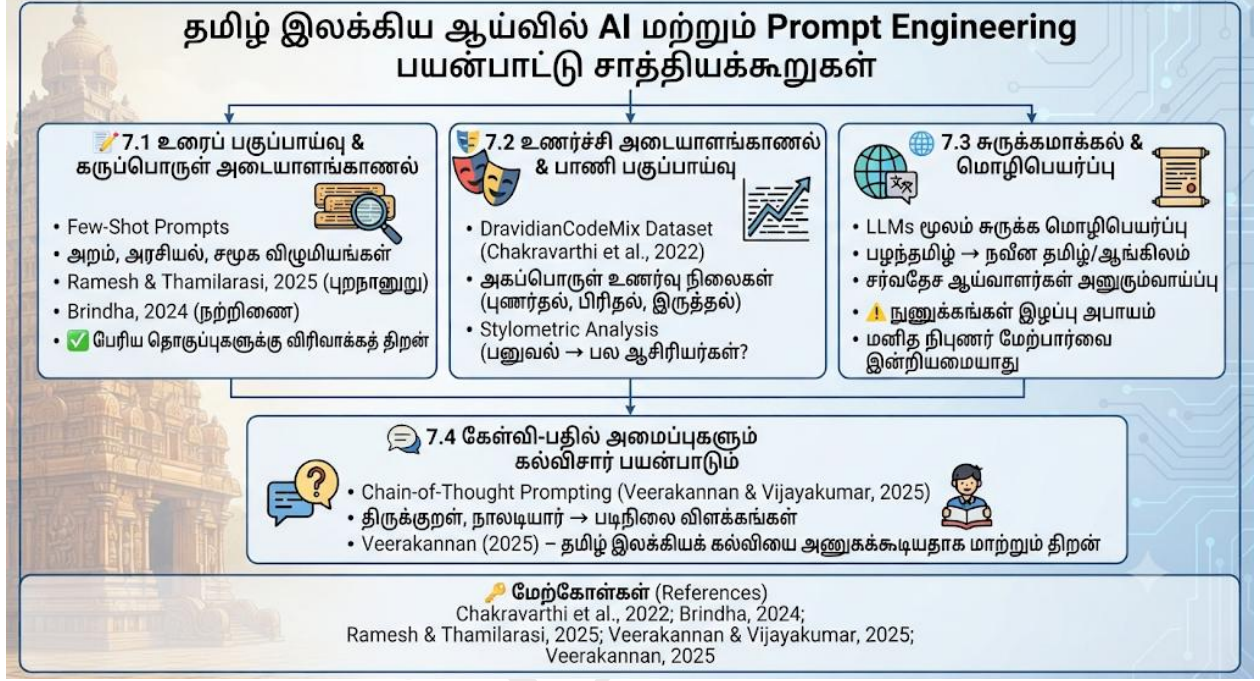


Fig5 : தமிழ் இலக்கிய ஆய்வில் AI மற்றும் Prompt Engineering பயன்பாட்டு

சாத்தியக்கூறுகள்

இப்படம் தமிழ் இலக்கிய ஆய்வுகளில் செயற்கை நுண்ணறிவு மற்றும் ப்ராம்ப்ட் இன்ஜினியரிங் எவ்வாறு பயன்படுகிறது என்பதை விளக்குகிறது. உரைப் பகுப்பாய்வு, உணர்ச்சி மற்றும் பாணி அடையாளங்காணல், சுருக்கமாக்கல், மொழிபெயர்ப்பு, மற்றும் கேள்வி-பதில் அமைப்புகள் மூலம் சங்க இலக்கியம் மற்றும் பழந்தமிழ் நூல்களை நவீன முறையில் ஆய்வு செய்ய இது உதவுகிறது. மேலும், சர்வதேச ஆய்வாளர்களுக்கான கல்விசார் பயன்பாடுகளையும் இதன் சாத்தியக்கூறுகளையும் இது தெளிவுபடுத்துகிறது.

8. விளக்கமளிக்கும் வழக்காடு: பழந்தமிழ் இலக்கியப் பகுப்பாய்வுக்கு Prompt Engineering

எடுத்துக்காட்டாக, ஒரு ஆய்வாளர் திருக்குறளின் அறத்துப்பாலில் இடம்பெறும் நல்லாட்சிக் கருத்துக்களை பகுப்பாய்வு செய்ய விரும்பினால், பின்வரும் மூன்று கட்ட Prompt Engineering அணுகுமுறையைப் பயன்படுத்தலாம்: முதலில், மாதிரிக்கு “தமிழ் இலக்கிய ஆய்வாளர்” எனும் பங்கை வழங்கி (Role-Based Prompting), இரண்டாவதாக, திருக்குறளின் தேர்ந்தெடுக்கப்பட்ட

குறள்களை Few-Shot எடுத்துக்காட்டுகளாக அளித்து, மூன்றாவதாக, Chain-of-Thought அணுகுமுறையில் ஒவ்வொரு குறளின் அரசியல்-நீதி கருத்தையும் படிநிலையாக விளக்கும்படி வழிமுறையிடலாம். இத்தகைய கட்டமைக்கப்பட்ட அணுகுமுறை, மாதிரியிடமிருந்து பொருள் இழப்பு குறைவான, சூழல் பொருத்தமான பகுப்பாய்வைப் பெற உதவுகிறது. எனினும், இறுதி முடிவுகள் எப்போதும் நிபுணத்துவம் வாய்ந்த மனித ஆய்வாளரால் சரிபார்க்கப்பட வேண்டும் ஏனெனில் மாதிரிகள் தவறான அல்லது கற்பனையான தகவல்களை உறுதியான தொனியில் வழங்கும் “மாயத் தோற்றம்” (Hallucination) எனும் பிரச்சினையைக் கொண்டுள்ளன (Sahoo et al., 2024).

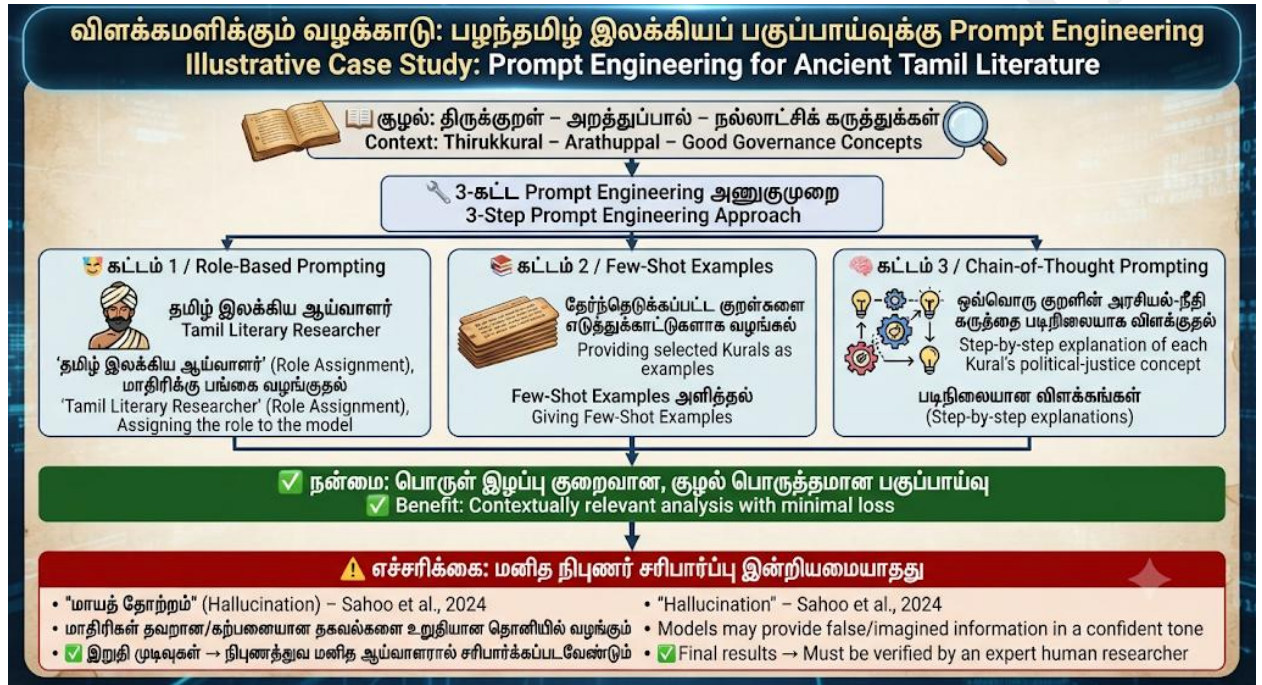


Fig 6: விளக்கமளிக்கும் வழக்காடு: பழந்தமிழ் இலக்கியப் பகுப்பாய்வுக்கு Prompt Engineering

இப்படம் திருக்குறள் அறத்துப்பாலின் நல்லாட்சிக் கருத்துக்களை பகுப்பாய்வு செய்ய மூன்று கட்ட ப்ராம்ப்ட் இன்ஜினியரிங் முறையைப் பயன்படுத்துவதை விளக்குகிறது. பங்கு-அடிப்படையிலான, சில உதாரண மற்றும் படிநிலைச் சிந்தனை அணுகுமுறைகள் மூலம் சூழல் பொருத்தமான பகுப்பாய்வு பெறப்படுகிறது. எனினும், மாதிரிகளின் மாயத் தோற்றச் சவாலைத் தவிர்க்க மனித நிபுணரின் சரிபார்ப்பு இன்றியமையாதது என்பதை இது எச்சரிக்கிறது.

9. சவால்களும் வரம்புகளும்

குறைவளமொழி நிலை தமிழுக்கான பயிற்சித் தரவு அளவு, ஆங்கிலம் போன்ற உயர்வளமொழிகளுடன் ஒப்பிடும்போது கணிசமாகக் குறைவானது (Varsha et al., 2025); இது

மாதிரிகளின் தமிழ் இலக்கண மற்றும் இலக்கியப் புரிதலைக் கட்டுப்படுத்துகிறது.

கலாச்சார நுணுக்கங்களின் இழப்பு பெருமொழி மாதிரிகள் பெரும்பாலும் ஆங்கில மொழிபெயர்ப்புகளை அடிப்படையாகக் கொண்ட தரவுகளில் பயிற்சி பெறுவதால், தமிழ் இலக்கியத்தின் தனித்துவமான கலாச்சார மற்றும் மொழியியல் நுணுக்கங்களைப் புரிந்துகொள்வதில் வரம்புகளைக் கொண்டுள்ளன (Varsha et al., 2025).

மாயத் தோற்றப் பிரச்சினை (Hallucination) மாதிரிகள் சில நேரங்களில் தவறான மேற்கோள்களையோ, இல்லாத பாடல் வரிகளையோ உறுதியான தொனியில் உருவாக்கக்கூடும்; இது இலக்கிய ஆய்வில் கடுமையான தவறான தகவல் பரவலுக்கு வழிவகுக்கும்.

கல்விசார் நேர்மை தொடர்பான கவலைகள் மாணவர்கள் மற்றும் ஆய்வாளர்கள் AI கருவிகளை அதிகமாகச் சார்ந்திருப்பது, சுயமான விமர்சன சிந்தனையையும் மூலநூல் நேரடி வாசிப்பையும் பாதிக்கக்கூடும்.

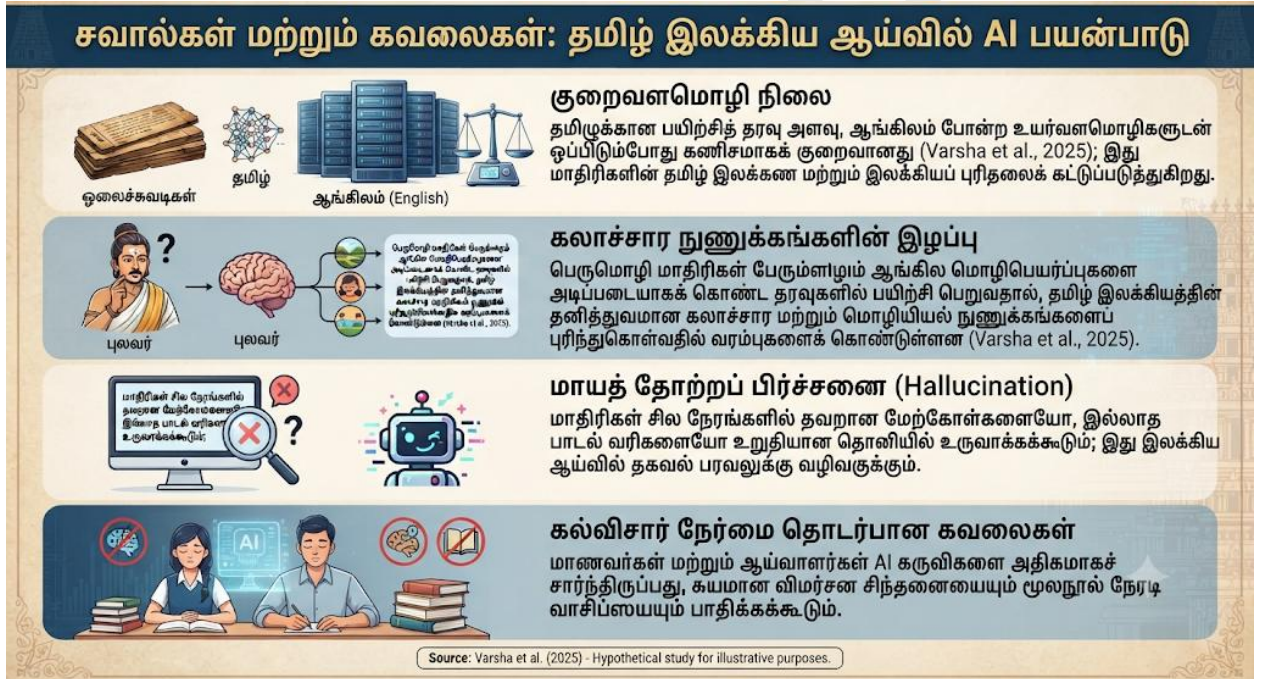


Fig 7: சவால்கள் மற்றும் கவலைகள்: தமிழ் இலக்கிய ஆய்வில் AI பயன்பாடு

இப்படம் தமிழ் இலக்கிய ஆய்வில் செயற்கை நுண்ணறிவைப் பயன்படுத்துவதன் முக்கிய சவால்களை விளக்குகிறது. தமிழுக்கான குறைந்த பயிற்சித் தரவு அளவு, கலாச்சார நுணுக்கங்களை இழத்தல் மற்றும் தவறான தகவல்களை வழங்கும் மாயத் தோற்றப் பிரச்சினை ஆகியவற்றுடன், மாணவர்களின் சுயமான விமர்சன சிந்தனையைப் பாதிக்கும் கல்விசார் நேர்மை கவலைகளையும் இது தெளிவாக எடுத்துரைக்கிறது.

10. நெறிமுறைக் கவலைகள்

தமிழ் இலக்கிய ஆய்வில் AI கருவிகளைப் பயன்படுத்தும்போது, மேற்கோள் துல்லியத்தையும் மூல ஆதாரங்களை சரிபார்க்கும் பொறுப்பையும் ஆய்வாளரே ஏற்றுக்கொள்ள வேண்டும். மாதிரி உருவாக்கும் விளக்கங்களை இறுதி முடிவுகளாக ஏற்காமல், அவற்றை கூடுதல் கருதுகோள்களாக மட்டுமே பயன்படுத்தி, பாரம்பரிய உரையாய்வு மற்றும் நிபுணர் மேற்பார்வையுடன் ஒப்பிட்டு உறுதிசெய்வது இன்றியமையாதது. மேலும், பயிற்சித் தரவுகளில் பயன்படுத்தப்படும் தமிழ் இலக்கியப் பனுவல்களின் பதிப்புரிமை மற்றும் பண்பாட்டு உரிமை தொடர்பான கேள்விகளும் கவனத்தில் கொள்ளப்பட வேண்டியவை.

11. எதிர்கால ஆராய்ச்சி திசைகள்

தமிழ் இலக்கிய ஆய்வுக்கென தனிப்பயனாக்கப்பட்ட, நிபுணர் சரிபார்த்த மேற்கோள் தரவுத்தொகுப்புகளை உருவாக்குதல் ஒரு முன்னுரிமை பணியாகும். மேலும், தமிழ் இலக்கிய கல்வெட்டியல், சுவடியியல் ஆகியவற்றுடன் இணைந்து செயல்படும் Optical Character Recognition (OCR) தொழில்நுட்பங்களை மேம்படுத்துவதன் மூலம், பாதுகாக்கப்படாத ஓலைச்சுவடி நூல்களை டிஜிட்டல் மயமாக்கி பகுப்பாய்வுக்கு உட்படுத்தும் வாய்ப்பும் உள்ளது. தமிழ் இலக்கிய ஆய்வாளர்களும், கணினி விஞ்ஞானிகளும் இணைந்து செயல்படும் பன்முகத் துறை ஆராய்ச்சிக் குழுக்களை உருவாக்குவது, இத்தொழில்நுட்பங்களை பொறுப்புடனும் துல்லியத்துடனும் பயன்படுத்துவதற்கு இன்றியமையாததாகும் (Veerakannan & Vijayakumar, 2025).

12. முடிவுரை

செயற்கை நுண்ணறிவும் Prompt Engineering நுட்பங்களும், தமிழ் இலக்கிய ஆய்வுக்கு பாரம்பரிய அணுகுமுறைகளை மாற்றுவதற்குப் பதிலாக, அவற்றை வலுப்படுத்தும் கருவிகளாக செயல்படும் திறனைக் கொண்டுள்ளன. Tamil-LLaMA போன்ற தமிழ்-குறிப்பிட்ட மொழி மாதிரிகளும், DravidianCodeMix போன்ற தரவுத்தொகுப்புகளும், தமிழ் மொழிக்கான கணினித் தமிழியல் உள்கட்டமைப்பை படிப்படியாக வலுப்படுத்தி வருகின்றன. எனினும், குறைவளமொழி வரம்புகள், கலாச்சார நுணுக்க இழப்பு, மாயத் தோற்றப் பிரச்சினை ஆகியவை, இத்தொழில்நுட்பங்களை எச்சரிக்கையுடன், மனித நிபுணத்துவத்துடன் இணைந்தே பயன்படுத்த வேண்டிய அவசியத்தை உணர்த்துகின்றன. இறுதியில், தமிழ் இலக்கிய ஆய்வாளர்களும் கணினி விஞ்ஞானிகளும் இணைந்து செயல்படும் பன்முகத் துறை ஒத்துழைப்பே, இத்தொழில்நுட்பங்களின் முழுத் திறனையும் பொறுப்புடன் பயன்படுத்திக்கொள்ள வழிவகுக்கும் என்பதில் ஐயமில்லை.

துணைநூற் பட்டியல் (References)

- [1.] Balachandran, A. (2023). Tamil-LLaMA: A new Tamil language model based on Llama 2. arXiv preprint arXiv:2311.05845.
- [2.] Brindha, S. (2024). நற்றிணை – மருதத்திணையில் நகை மெய்ப்பாடு. International Journal of Creative Research Thoughts (IJCRT), 12(11), a611–a618.
- [3.] Chakravarthi, B. R., Priyadharshini, R., Muralidaran, V., Jose, N., Suryawanshi, S., Sherly, E., & McCrae, J. P. (2022). DravidianCodeMix: Sentiment analysis and offensive language identification dataset for Dravidian languages in code-mixed text. Language Resources and Evaluation, 56(3), 765–806.
- [4.] Doddapaneni, S., Aralikatte, R., Ramesh, G., Goyal, S., Khapra, M. M., Kunchukuttan, A., & Kumar, P. (2023). Towards leaving no Indic language behind: Building monolingual corpora, benchmark and models for Indic languages. Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, 1, 12402–12426.
- [5.] Minaee, S., Mikolov, T., Nikzad, N., Chenaghlu, M., Socher, R., Amatriain, X., & Gao, J. (2024). Large language models: A survey. arXiv preprint arXiv:2402.06196.
- [6.] Premjith, B., Bharathi, B., Chakravarthi, B. R., et al. (2023). Findings of the shared task on multimodal sentiment analysis and emotion analysis in Tamil and Malayalam. Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages.
- [7.] Raja, R. (2022). சங்க கால வாழ்வியல் விழுமியங்கள்: புறநானூற்றில் அரசியல் அறங்கள். International Journal of Creative Research Thoughts (IJCRT), 10(3), 849–855.
- [8.] Ramesh, T. K., & Thamilarasi, P. (2025). Sense of self & related behaviour: A behavioral psychology perspective in Purananooru. International Journal of Creative Research Thoughts (IJCRT), 13(8), g173–g179.
- [9.] Ranathunga, S., Lee, E.-S. A., Prifti Skenduli, M., Shekhar, R., Alam, M., & Kaur, R. (2023). Neural machine translation for low-resource languages: A survey. ACM Computing Surveys, 55(11), 1–37.
- [10.] Sahoo, P., Singh, A. K., Saha, S., Jain, V., Mondal, S., & Chadha, A. (2024). A systematic survey of prompt engineering in large language models: Techniques and applications. arXiv preprint arXiv:2402.07927.

- [11.] Sopa, R., & Murugeswari, N. (2023). பதினெண்கீழ்க்கணக்கு நூல்களில் கற்பு. *International Journal of Creative Research Thoughts (IJCRT)*, 11(3), 745–750.
- [12.] Varsha, J., et al. (2025). From phonemes to meaning: Evaluating large language models on Tamil. *arXiv preprint arXiv:2511.12387*.
- [13.] Vatsal, S., & Dubey, H. (2024). A survey of prompt engineering methods in large language models for different NLP tasks. *arXiv preprint arXiv:2407.12994*.
- [14.] Veerakannan, S. (2025). Artificial intelligence technology: A boon in writing Tamil essays. *Tamilmanam International Research Journal of Tamil Studies*, 1(04), 221–228. <https://doi.org/10.63300/dn796x92>
- [15.] Veerakannan, S., & Vijayakumar, M. (2025). Artificial intelligence for Tamil literature: An overview. *Tamilmanam International Research Journal of Tamil Studies*, 1(07), 363–370. <https://doi.org/10.63300/eanb2p66>