தமிழ் செயற்கை நுண்ணறிவு ஆய்விதழ்
மதிப்பாய்வு செய்யப்படும் இதழ்

Tamil Journal of Artificial Intelligence
Peer-Reviewed Journal

Volume 1, Issue 1 – January 2026

தமிழ் ஓலைச்சுவடிகளைப் பாதுகாக்கச் செயற்கை நுண்ணறிவின் (AI) பயன்பாடு

முனைவர் த. மகேஸ்வரி, முதன்மை ஆசிரியர், தமிழ்மொழி மற்றும் இலக்கிய பன்னாட்டு ஆய்விதழ் (IJTLLS), விருதுநகர், தமிழ்நாடு, இந்தியா.
ORCiD: https://orcid.org/0000-0002-4187-0120

அறிமுகம்

தமிழ் நாகரிகத்தின் தொன்மையான அடையாளமாக விளங்கும் ஓலைச்சுவடிகள், நம் முன்னோர்களின் அறிவுச் செல்வத்தை இன்றளவும் வெளிப்படுத்தும் ஆதாரமாகத் திகழ்கின்றன. சங்க இலக்கியங்கள் முதல் சித்தர் பாடல்கள் வரை, உளவியல், மருத்துவம், வானியல் போன்ற பல்வேறு அறிவியல் துறைகளின் முதன்மை ஆதாரங்களாக இவை விளங்குகின்றன. எனினும், காலத்தின் தொடர்ச்சியான சீற்றத்துக்கு இலக்காகி, இந்த அரிய அறிவுப் பொக்கிஷங்கள் மெல்ல மெல்ல சிதைந்து மறையும் அபாயத்தில் உள்ளன. இந்த நெருக்கடியான சூழலில், செயற்கை நுண்ணறிவு (AI) தொழில்நுட்பமானது ஓலைச்சுவடிகளைப் பாதுகாக்கும் பணியில் ஒரு புதிய நம்பிக்கை தருவதாக அமைந்து உள்ளது. பழங்கால ஓலைச்சுவடிகளை எண்ணிம (Digital) முறையில் பாதுகாப்பதில் செயற்கை நுண்ணறிவு மாற்றங்களை ஏற்படுத்துகிறது என்பதைப் பற்றி இக்கட்டுரை விளக்குகிறது.

ஓலைச்சுவடிகளைப் பாதுகாப்பதில் உள்ள சவால்கள்

ஓலைச்சுவடிகளைப் பாதுகாக்கும் பணியில் நாம் பலவிதமான சவால்களை எதிர்கொள்ள வேண்டியுள்ளது. முதலாவதாக, ஓலைச்சுவடிகள் பனை இழைகளால் ஆனவை என்பதால், ஈரப்பதம், தட்பவெப்ப நிலை மாற்றங்கள் மற்றும் தொடர்ச்சியான பயன்பாட்டினால் ஏற்படும் தேய்மானம் ஆகியவற்றால் அவை எளிதில் சிதைந்துவிடுகின்றன. குறிப்பாக, ஆயிரம் முதல் பல நூறு ஆண்டுகளுக்கு முற்பட்ட பழமையான ஓலைகள் சிதைந்த நிலையில் காணப்படுகின்றன. இரண்டாவதாக, கரப்பான் பூச்சி (Cockroach), சில்வர் ஃபிஷ் (Silverfish) போன்ற பூச்சிகளின் தாக்குதலால் பல அரிய சுவடிகள் நிரந்தரமாக அழிந்துபோன நிலையில் உள்ளன. இத்தகைய சிதைவுகளுக்கு அப்பால், சுவடிகளில் உள்ள எழுத்து வடிவங்களும் படிப்பதற்கு பெரும் சவாலாக அமைகின்றன. பண்டைய தமிழ் எழுத்து முறைகள், கிரந்த எழுத்துகள், வட்டெழுத்துகள் மற்றும் சுவடிகளுக்கே உரிய தனித்துவமான சுருக்கெழுத்துகள், குறியீடுகள் போன்றவை தற்கால செயற்கை நுண்ணறிவு மூலம் 'படிம எழுத்துணரி' (OCR) தொழில்நுட்பம் செயலாக்கம் செய்யப்படலாம். மேலும், சங்ககால மற்றும் இடைக்காலத் தமிழ் மொழியின் சொல்லாட்சிகளும் இலக்கண அமைப்புகளும் நவீனத் தமிழிலிருந்து பெரிதும் வேறுபடுவதால், செயற்கை நுண்ணறிவு அமைப்புகளுக்கு இவற்றைப் புரிந்துகொள்ளக் கூடுதல் தரவுகளும் பயிற்சியும் தேவைப்படுகின்றன.

எண்ணிம மயமாக்கலில் செயற்கை நுண்ணறிவின் பங்கு

ஓலைச்சுவடிகளைப் பாதுகாப்பாக எண்ணிம (Digital) வடிவத்திற்கு மாற்றுவதில் பல்வேறு செயற்கை நுண்ணறிவுத் தொழில்நுட்பங்கள் ஒருங்கிணைந்து செயல்படுத்தலாம். இதன் முதல் படியாக இருப்பது உயர்தர ஸ்கேனிங் மற்றும் படமாக்கல் முறையாகும். மல்டி-ஸ்பெக்ட்ரல் (Multi-spectral) இமேஜிங் தொழில்நுட்பத்தின் மூலம் பல்வேறு ஒளி அலைநீளங்களைப் பயன்படுத்தி, கண்ணுக்குத் தெரியாமல் மறைந்துபோன ஓலைச் சுவடிஎழுத்துகளைக் கண்டறிய முடிகிறது. மேலும், முப்பரிமாண (3D) ஸ்கேனிங் மற்றும் 1200 DPI-க்கும் அதிகமான துல்லியத்தன்மை கொண்ட படமாக்கல் முறை மூலம் ஓலையின்

Volume 1, Issue 1 – January 2026

அமைப்பு மற்றும் சிதைவுகள் மிகத் தெளிவாகப் பதிவு செய்ய முடியும். இதனைத் தொடர்ந்து, படமாக்கப்பட்ட சுவடிகளிலிருந்து எழுத்துகளைப் பிரித்தெடுக்க கணினியின் துணைகொண்டு 'ஆழமான கற்றல்' **(Deep Learning)** மூலம் அடிப்படையிலான பிரத்தியேகத் தமிழ் 'ஓ.சி.ஆர்' **(OCR)** அமைப்புகள் பயன்படுத்தலாம். இவை பழைய எழுத்து வடிவங்களை மிகத் துல்லியமாக அடையாளம் கண்டு செயற்கை நுண்ணறிவு நெறிமுறைகள் மூலம் பிழைகளைத் தானாகவே சரிசெய்துகொள்கின்றன. இவ்வாறு ஓலைச்சுவடிகளில் இருந்து பிரித்தெடுக்கப்பட்ட பண்டைய தமிழ் உரைகளைப் புரிந்துகொள்வதற்கு 'இயற்கை மொழிச் செயலாக்க' **(NLP)** மாதிரிகள் பெருமளவில் உதவுகின்றன. சங்க இலக்கியங்கள் மற்றும் இடைக்கால நூல்களுக்கெனச் சிறப்பாகப் பயிற்றுவிக்கப்பட்ட இந்த மாதிரிகள், வாக்கியங்களின் வரலாற்றுச் சூழல் மற்றும் சொற்பொருளைப் பகுப்பாய்வு செய்து, பழைய தமிழை நவீன தமிழுக்குத் தானியங்கியாக மொழிபெயர்க்கும் பணியையும் செவ்வனே செய்கின்றன. எல்லாவற்றிற்கும் மேலாக, சிதைந்த ஓலைகளை எண்ணிம முறையில் மறுசீரமைப்பதில் செயற்கை நுண்ணறிவின் பங்கு அளப்பரியது எனலாம். ஓலைச் சுவடிகளில் அழிந்துபோன பகுதிகளை நிரப்புதல், மங்கிப்போன எழுத்தை மீட்டெடுத்தல் மற்றும் காணாமல் போன வரிகளை மொழிநடைக்கேற்ப ஊகித்து முழுமைப்படுத்துதல் போன்ற வியக்கத்தக்க பணிகளை செயற்கை நுண்ணறிவு சார்ந்த **OCR** தொழில்நுட்பம் சாத்தியமாக்குகிறது.

ஓலைச்சுவடி பாதுகாப்புக்கான படிநிலைச் செயல்முறை

ஓலைச்சுவடிகளைப் பாதுகாக்கும் இந்தச் செயற்கை நுண்ணறிவுச் செயல்முறை பல முக்கியப் படிநிலைகளைக் கொண்டுள்ளது. தொடக்கக் கட்டமாக, ஓலைகளை மென்மையாக்குதல், ஈரப்பதம் மற்றும் வெப்பநிலையைக் கட்டுப்படுத்துதல், பூச்சி மற்றும் தூசிகளை அகற்றுதல் போன்ற முன்-செயலாக்கப் **(Pre-processing)** பணிகள் மேற்கொள்ள வேண்டும். அதன் பின்னர், பல கோணங்களில் முப்பரிமாண ஸ்கேனர்கள் **(3 D Scanners)** மற்றும் மல்டி-ஸ்பெக்ட்ரல் கேமராக்கள் **(Multi Spectral Cameras)** மூலம் ஒளி, நிழல் சமநிலைப்படுத்தப்பட்டு உயர்தரப் படங்களாக அவை மாற்றம் செய்யப்பட வேண்டும். அடுத்த கட்டமாக, பிரத்தியேக மென்பொருள்கள் மூலம் எழுத்துகள் பிரித்தெடுக்கப்பட்டு, பிழை திருத்தங்கள் மேற்கொள்ளப்பட வேண்டும். அதைத் தொடர்ந்து, இயற்கை மொழிச் செயலாக்கக் **(NLP)** கருவிகளின் உதவியுடன் அந்த உரையின் சொற்பொருள், இலக்கணம் மற்றும் வரலாற்றுச் சூழல் ஆகியவற்றை ஆழமாகப் பகுப்பாய்வு செய்யப்பட வேண்டும். இறுதியாக, இவ்வாறு பெறப்பட்ட தகவல்கள் அனைத்தும் பன்மொழி இடைமுகங்களைக் கொண்ட, எளிதில் தேடிக் கண்டறியக்கூடிய எண்ணிமத் தரவுத்தளங்களில் ஆராய்ச்சியாளர்களுக்கும் பொதுப் பயன்பாட்டிற்காக காப்பகப்படுத்த வேண்டும்.

முடிவுரை

ஓலைச்சுவடிகளைப் பாதுகாக்கச் செயற்கை நுண்ணறிவுத் தொழில்நுட்பத்தைப் பயன்படுத்துவது என்பது ஒரு தொழில்நுட்ப முன்னேற்றம் மற்றுமல்லாமல்; அது தமிழ் கலாச்சாரப் பாரம்பரியத்தைப் பாதுகாப்பதற்கான ஒரு வரலாற்றுச் சிறப்புமிக்கப் செயலாகும் எனலாம். ஓ.சி.ஆர் **(OCR)** தொழில்நுட்பம், இயற்கை மொழிச் செயலாக்கம் **(NLP)** மற்றும் முப்பரிமாண ஸ்கேனிங் **(3 D Scanning)** போன்ற நவீனத் தொழில்நுட்பங்களைப் பயன்படுத்தி சிதைந்துபோகும் நிலையிலுள்ள பனை ஓலைச் சுவடிகளில் உள்ள அரிய தகவல்களை எக்காலத்திற்கும் அழியாமல் பாதுகாக்க முடியும். எனவே, ஒருங்கிணைந்த எண்ணிம வடிவமாக்கும் முயற்சிகள், தமிழ் மொழியின் தொன்மையையும் தமிழர்களின் இணையற்ற கலாச்சாரத்தின் எதிர்காலத்தையும் பாதுகாக்கும் அரணாக விளங்கும் என்பது திண்ணம் .

Volume 1, Issue 1 – January 2026

ஆகையால், நமது எதிர்கால சந்ததிகள் இவைபோன்ற கணினி துறைகளை செவ்வன கற்று ஆவனப்படுத்தும் முறைகளையும் கற்றுக்கொள்ள வேண்டும்.

குறிப்புகள்

[1] கார்த்திகேயன், சு., & அருள், அ. (2023). சித்தாந்த ஓலைச்சுவடிகளை மீட்டமைக்கப் படச் செயலாக்கம் (Image Processing) மற்றும் செயற்கை நுண்ணறிவின் பங்கு. இந்தியத் தகவல் தொழில்நுட்ப ஆய்விதழ், 19(1), 55-68.
[2] சிவசந்த், மு., & கபிலன், வ. (2024). டிரான்ஸ்பார்மார் (Transformer) தொழில்நுட்பம் வழியாகத் தானியங்கியான தமிழ் ஓலைச்சுவடிகளை அறிதலும் பாதுகாத்தலும். செயற்கை நுண்ணறிவு மற்றும் தரவு அறிவியல் ஆய்விதழ், 15(2), 45-58.
[3] பார்த்திபன், க., & முருகன், தவ. (2022). ஓலைச்சுவடிகளில் உள்ள தமிழ் எழுத்துக்களைத் தானியங்கி முறையில் கண்டறிவதற்கான கன்வல்யூஷனல் நியூரல் நெட்வார்க் (CNN) அணுகுமுறை. பன்னாட்டு கணினியியல் ஆய்விதழ், 10(4), 200-215.
[4] பவல்முருகன், கு. (2023). தமிழ் ஓலைச்சுவடி எழுத்துக்களை அறிவதற்கான ஆழ் கற்றல் (Deep Learning) வழிமுறைகள் மற்றும் அதன் பயன்பாடுகள். தமிழ் கணினி தொழில்நுட்ப மாநாட்டு ஆய்வுக் கோவை, 8, 112-120.
[5] AI4Bharat. (பதவி இல்லை). முகப்புப் பக்கம். https://ai4bharat.iitm.ac.in/
[6] Institut Français de Pondichéry (IFP). (பதவி இல்லை). முகப்புப் பக்கம். https://www.ifpindia.org/
[7] கிட்ஹப் (GitHub). (பதவி இல்லை). தமிழ் ஓலைச்சுவடி வாசிப்புத் திட்டங்கள் (Tamil Palm Leaf Character Recognition). https://github.com/
[8] தஞ்சாவூர் மகாராஜா சரபோஜியின் சாரஸ்வதி மகால் நூலகம். (பதவி இல்லை). முகப்புப் பக்கம். https://sarasvathimahallibrary.tn.gov.in/
[9] தமிழ் இணையக் கல்விக் கழகம் (Tamil Virtual Academy). (பதவி இல்லை). முகப்புப் பக்கம். https://www.tamilvu.org/

நிதிசார் கட்டுரையாளர் உறுதிமொழி: இல்லை
கட்டுரையாளர் நன்றியுரை: இல்லை
கட்டுரையாளர் உறுதிமொழி: இக்கட்டுரையில் எவ்வித முரண்பாடும் இல்லை என்று உறுதிமொழி அளிக்கிறேன்.

இக்கட்டுரை கிரியேட்டிவ் காமன்சு ஆட்ரிபியூசன் 4.0வின் **Creative Commons Attribution4.0** கீழ் பன்னாட்டு உரிமம் பெற்றுள்ளது.