ಭಾಷಾ ತಂತ್ರಜ್ಞಾನ ಮತ್ತು ಕೃತಕ ಬುದ್ಧಿಮತ್ತೆ: ಒಂದು ಸಮಗ್ರ ಅವಲೋಕನ

ಚೇತನ್ ಜಿ.ಜೆ.[1], ತೇಜಸ್ ಶೆಟ್ಟಿ ಜಿ.[2] & ಮನೋಜ್ ಎಂ.[3]

[1]ದ್ವಿತೀಯ ಬಿ.ಕಾಂ., ಸಂತ ಫ್ರಾನ್ಸಿಸ್ ಕಾಲೇಜು, ಕೋರಮಂಗಲ, ಬೆಂಗಳೂರು.

[2]ದ್ವಿತೀಯ ಬಿ.ಕಾಂ., ಸಂತ ಫ್ರಾನ್ಸಿಸ್ ಕಾಲೇಜು, ಕೋರಮಂಗಲ, ಬೆಂಗಳೂರು.

[3]ಸಹಾಯಕ ಪ್ರಾಧ್ಯಾಪಕರು, ಸಂತ ಫ್ರಾನ್ಸಿಸ್ ಕಾಲೇಜು, ಕೋರಮಂಗಲ, ಬೆಂಗಳೂರು.

DOI: https://doi.org/10.5281/zenodo.18367784

ABSTRACT:

ಪ್ರಸ್ತುತ ಸಂಶೋಧನಾ ಲೇಖನವು ಭಾಷಾ ತಂತ್ರಜ್ಞಾನ ಮತ್ತು ಕೃತಕ ಬುದ್ಧಿಮತ್ತೆಯ (AI) ಸಮ್ಮಿಲನವನ್ನು, ವಿಶೇಷವಾಗಿ ಕನ್ನಡ ಭಾಷೆಯ ಹಿನ್ನೆಲೆಯಲ್ಲಿ ಸಮಗ್ರವಾಗಿ ಪರಿಶೀಲಿಸುತ್ತದೆ. ಸಾಂಕೇತಿಕ ವಿಧಾನಗಳಿಂದ ಆಧುನಿಕ ಡೀಪ್ ಲರ್ನಿಂಗ್ ಮತ್ತು ಟ್ರಾನ್ಸ್‌ಫಾರ್ಮರ್ ಮಾದರಿಗಳವರೆಗಿನ ನೈಸರ್ಗಿಕ ಭಾಷಾ ಸಂಸ್ಕರಣೆಯ (NLP) ವಿಕಸನವನ್ನು ಇದು ಚರ್ಚಿಸುತ್ತದೆ. ಕನ್ನಡದಂತಹ ಸಂಕೀರ್ಣ ವ್ಯಾಕರಣ ಮತ್ತು ಪ್ರತ್ಯಯಾಧಾರಿತ (agglutinative) ಭಾಷೆಗಳು ಎದುರಿಸುತ್ತಿರುವ ದತ್ತಾಂಶದ ಕೊರತೆ, ಉಪಭಾಷೆಗಳ ವೈವಿಧ್ಯತೆ ಮತ್ತು ತಾಂತ್ರಿಕ ಸವಾಲುಗಳನ್ನು ಲೇಖನವು ವಿಶ್ಲೇಷಿಸುತ್ತದೆ. ಓಸಿಆರ್ (OCR), ಯೂನಿಕೋಡ್ ಮತ್ತು ಧ್ವನಿ ಆಧಾರಿತ ತಂತ್ರಜ್ಞಾನಗಳ ಪ್ರಾಯೋಗಿಕ ಅನ್ವಯಗಳನ್ನು ವಿವರಿಸುತ್ತಾ, ಜ್ಞಾನದ ಜನತಂತ್ರೀಕರಣದಲ್ಲಿ AI ಪಾತ್ರವನ್ನು ಗುರುತಿಸಲಾಗಿದೆ. ಅಂತಿಮವಾಗಿ, ಡಿಜಿಟಲ್ ಯುಗದಲ್ಲಿ ಭಾಷಾ ಅಸ್ಮಿತೆಯನ್ನು ಉಳಿಸಿಕೊಂಡು ತಾಂತ್ರಿಕ ಪ್ರಗತಿಯನ್ನು ಸಾಧಿಸುವ ಅಗತ್ಯವನ್ನು ಈ ಅಧ್ಯಯನ ಪ್ರತಿಪಾದಿಸುತ್ತದೆ.

KEYWORDS:

ಭಾಷಾ ತಂತ್ರಜ್ಞಾನ, ಕೃತಕ ಬುದ್ಧಿಮತ್ತೆ, ನೈಸರ್ಗಿಕ ಭಾಷಾ ಸಂಸ್ಕರಣೆ, ಯಂತ್ರ ಕಲಿಕೆ, ಕನ್ನಡ ಕಂಪ್ಯೂಟಿಂಗ್.

ಪೀಠಿಕೆ:

ಭಾಷೆಯು ಮಾನವಕುಲದ ಸಂವಹನದ ಮೂಲಾಧಾರ ಮತ್ತು ಜ್ಞಾನದ ವಾಹಕವಾಗಿದೆ. ಇಂದಿನ ಡಿಜಿಟಲ್ ಯುಗದಲ್ಲಿ, ಮಾನವನ ಭಾಷೆಯನ್ನು ಯಂತ್ರಗಳು ಅರ್ಥಮಾಡಿಕೊಳ್ಳುವಂತೆ ಮಾಡುವುದೇ 'ಭಾಷಾ ತಂತ್ರಜ್ಞಾನ'. ಈ ತಂತ್ರಜ್ಞಾನಕ್ಕೆ ಕೃತಕ ಬುದ್ಧಿಮತ್ತೆಯು (AI) ಜೀವ ತುಂಬುತ್ತಿದ್ದು, ಒಂದು ಹೊಸ ಕ್ರಾಂತಿಯನ್ನೇ ಸೃಷ್ಟಿಸಿದೆ. ಕಂಪ್ಯೂಟರ್‌ಗಳು ಮನುಷ್ಯರಂತೆ ಓದುವ, ಕೇಳುವ ಮತ್ತು ಮಾತನಾಡುವ ಸಾಮರ್ಥ್ಯ ಪಡೆಯುತ್ತಿವೆ.

ಇದರ ಪ್ರಮುಖ ಅಂಗವಾದ 'ನೈಸರ್ಗಿಕ ಭಾಷಾ ಸಂಸ್ಕರಣೆ' (NLP) ಪದಗಳ ಹಿಂದಿನ ಗಹನವಾದ ಅರ್ಥವನ್ನು ವಿಶ್ಲೇಷಿಸುತ್ತದೆ. ಹಿಂದೆ ಕೇವಲ ಗಣಿತದ ಸೂತ್ರಗಳಿಗೆ ಸೀಮಿತವಾಗಿದ್ದ ಯಂತ್ರಗಳು, ಇಂದು ಭಾವನೆಗಳನ್ನು ಗ್ರಹಿಸುವ ಮಟ್ಟಕ್ಕೆ ಬೆಳೆದಿವೆ. ಕೃತಕ ಬುದ್ಧಿಮತ್ತೆಯು ಕೋಟ್ಯಂತರ ದತ್ತಾಂಶಗಳ ಮೂಲಕ ಭಾಷೆಯ ಸೂಕ್ಷ್ಮ ವ್ಯಾಕರಣ ಮತ್ತು ಶೈಲಿಯನ್ನು ಕಲಿಯುತ್ತಿದೆ.

ಗೂಗಲ್ ಟ್ರಾನ್ಸ್‌ಲೇಟ್‌ನಂತಹ ಸಾಧನಗಳು ಇಂದು ಭಾಷೆಯ ಗಡಿಗಳನ್ನು ಮೀರಿ ಜಗತ್ತನ್ನು ಬೆಸೆಯುತ್ತಿವೆ. ಧ್ವನಿ ಆಧಾರಿತ ಅಲೆಕ್ಸಾ ಅಥವಾ ಸಿರಿಯಂತಹ ಸಹಾಯಕರು ನಮ್ಮ ದೈನಂದಿನ ಕೆಲಸಗಳನ್ನು ಸುಲಭಗೊಳಿಸುತ್ತಿದ್ದಾರೆ. ಆದರೆ, ಇಂಗ್ಲಿಷ್‌ಗೆ ಹೋಲಿಸಿದರೆ ಕನ್ನಡದಂತಹ ಪ್ರಾದೇಶಿಕ ಭಾಷೆಗಳಿಗೆ ದತ್ತಾಂಶದ ಕೊರತೆಯ ದೊಡ್ಡ ಸವಾಲಿದೆ. ನಮ್ಮ ಭಾಷೆಯ ಸಂಕೀರ್ಣ ವ್ಯಾಕರಣ ಮತ್ತು ವಿಭಕ್ತಿ ಪ್ರತ್ಯಯಗಳನ್ನು ಗ್ರಹಿಸುವುದು AIಗೆ ಇನ್ನೂ ಕಷ್ಟಕರವಾಗಿದೆ. ಪ್ರಾದೇಶಿಕ ಉಪಭಾಷೆಗಳ ವೈವಿಧ್ಯತೆಯು ತಂತ್ರಜ್ಞಾನದ ಅಳವಡಿಕೆಗೆ ಹೊಸ ಅಡೆತಡೆಗಳನ್ನು ಒಡ್ಡುತ್ತಿದೆ.

ಆದರೂ, ಡಿಜಿಟಲ್ ಲೋಕದಲ್ಲಿ ಕನ್ನಡವನ್ನು ಉಳಿಸಿ ಬೆಳೆಸಲು ಈ ತಂತ್ರಜ್ಞಾನವು ಅತ್ಯಂತ ಅನಿವಾರ್ಯವಾಗಿದೆ. ಶಿಕ್ಷಣ ಮತ್ತು ಆಡಳಿತವನ್ನು ಸಾಮಾನ್ಯ ಜನರಿಗೆ ಅವರದ್ದೇ ಭಾಷೆಯಲ್ಲಿ ತಲುಪಿಸಲು AI ದೊಡ್ಡ ಶಕ್ತಿಯಾಗಲಿದೆ. ಜಾಗತಿಕ ವೇದಿಕೆಯಲ್ಲಿ ಕನ್ನಡದ ಅಸ್ಮಿತೆಯನ್ನು ಕಾಪಾಡಲು ಭಾಷಾ ತಂತ್ರಜ್ಞಾನವು ಒಂದು ಸೇತುವೆಯಾಗಿ ಕಾರ್ಯನಿರ್ವಹಿಸುತ್ತದೆ. ಮುಂದಿನ ದಿನಗಳಲ್ಲಿ ಭಾಷಾ ತಡೆಗೋಡೆಗಳಿಲ್ಲದ ಜಗತ್ತನ್ನು ನಿರ್ಮಿಸುವುದು ಈ ತಂತ್ರಜ್ಞಾನದ ಪ್ರಮುಖ ಗುರಿಯಾಗಿದೆ. ತಂತ್ರಜ್ಞಾನ ಮತ್ತು ಸಂಸ್ಕೃತಿಯ ಈ ಸಮ್ಮಿಲನವು ಮಾನವ ಇತಿಹಾಸದ ಹೊಸ ಅಧ್ಯಾಯಕ್ಕೆ ನಾಂದಿ ಹಾಡಿದೆ.

ನೈಸರ್ಗಿಕ ಭಾಷಾ ಸಂಸ್ಕರಣೆ (NLP): ಸಂಕ್ಷಿಪ್ತ ವಿಶ್ಲೇಷಣೆ

ನೈಸರ್ಗಿಕ ಭಾಷಾ ಸಂಸ್ಕರಣೆಯು ಯಂತ್ರಗಳಿಗೆ ಮಾನವ ಭಾಷೆಯ ಸಂಕೀರ್ಣತೆಯನ್ನು ಅರ್ಥಮಾಡಿಸುವ ಒಂದು ಕಲೆ. ಇದರ ಪ್ರಮುಖ ಹಂತಗಳ ವಿಶ್ಲೇಷಣೆ ಇಲ್ಲಿದೆ:

ಪದಗಳ ವಿಭಜನೆ (Morphology): ಯಂತ್ರವು ಮೊದಲು ಪದಗಳನ್ನು ಅದರ ಮೂಲ ರೂಪಕ್ಕೆ ತರುತ್ತದೆ. ಕನ್ನಡದಂತಹ ಸಂಕೀರ್ಣ ಭಾಷೆಯಲ್ಲಿ 'ಮನೆಗೆ' ಎಂಬ ಪದವನ್ನು 'ಮನೆ + ಗೆ' ಎಂದು ಬಿಡಿಸಿ ನೋಡುವುದು ಅತ್ಯಗತ್ಯ. ಇಲ್ಲಿ ಪದದ ಮೂಲ ಧಾತು (Root) ಮತ್ತು ಪ್ರತ್ಯಯಗಳನ್ನು ಗುರುತಿಸುವುದರಿಂದ ವಾಕ್ಯದ ಉದ್ದೇಶ ತಿಳಿಯುತ್ತದೆ.

ವಾಕ್ಯದ ಚೌಕಟ್ಟು (Syntax): ಇದು ಭಾಷೆಯ ವ್ಯಾಕರಣದ ರಚನೆಯನ್ನು ವಿಶ್ಲೇಷಿಸುತ್ತದೆ. 'Parsing' ಎಂಬ ಪ್ರಕ್ರಿಯೆಯ ಮೂಲಕ ಕರ್ತೃ, ಕರ್ಮ ಮತ್ತು ಕ್ರಿಯಾಪದಗಳ ನಡುವಿನ ಸಂಬಂಧವನ್ನು ಯಂತ್ರ ಪತ್ತೆ ಮಾಡುತ್ತದೆ. ಇಂಗ್ಲಿಷ್ (SVO) ಮತ್ತು ಕನ್ನಡದ (SOV) ವಾಕ್ಯ ರಚನಾ ವ್ಯತ್ಯಾಸಗಳನ್ನು ಗುರುತಿಸುವುದು ಈ ಹಂತದ ಮುಖ್ಯ ಉದ್ದೇಶ.

ಅರ್ಥದ ಹುಡುಕಾಟ (Semantics): ಕೇವಲ ವ್ಯಾಕರಣ ಸರಿಯಿದ್ದರೆ ಸಾಲದು, ವಾಕ್ಯವು ಅರ್ಥಪೂರ್ಣವಾಗಿರಬೇಕು. ಒಂದೇ ಪದಕ್ಕೆ ಹಲವು ಅರ್ಥಗಳಿದ್ದಾಗ (ಉದಾ: ಹೊಳೆ), ಸಂದರ್ಭಕ್ಕೆ ತಕ್ಕಂತೆ ಸರಿಯಾದ ಅರ್ಥವನ್ನು 'Word Embeddings' ತಂತ್ರಜ್ಞಾನದ ಮೂಲಕ AI ಗ್ರಹಿಸುತ್ತದೆ. ಇದು ಭಾಷಾಂತರದಲ್ಲಿ ನಿಖರತೆಯನ್ನು ತರುತ್ತದೆ.

ಧ್ವನಿಯಿಂದ ಪಠ್ಯಕ್ಕೆ (Speech-to-Text): ಧ್ವನಿ ತರಂಗಗಳನ್ನು ಡಿಜಿಟಲ್ ರೂಪಕ್ಕೆ ತರುವ ಪ್ರಕ್ರಿಯೆ ಇದು. ಮಾತಿನ ಏರಿಳಿತಗಳನ್ನು ಅಕ್ಷರಗಳಾಗಿ (Phonemes) ಪರಿವರ್ತಿಸಿ, ಸುತ್ತಮುತ್ತಲಿನ ಶಬ್ದಗಳನ್ನು (Noise) ನಿವಾರಿಸಿ ಸ್ಪಷ್ಟವಾದ ಪಠ್ಯವನ್ನು ಸೃಜಿಸಲಾಗುತ್ತದೆ. ಇದಕ್ಕೆ ಇಂದು 'Deep Learning' ಮಾದರಿಗಳನ್ನು ಹೆಚ್ಚಾಗಿ ಬಳಸಲಾಗುತ್ತಿದೆ.

ಕನ್ನಡ ಭಾಷಾ ತಂತ್ರಜ್ಞಾನ: ಅನ್ವಯಿಕ ವಿಶ್ಲೇಷಣೆ

ಕನ್ನಡವು ಒಂದು ಅಂಟಿಸುವ ಭಾಷೆ (Agglutinative Language). ಅಂದರೆ ಒಂದು ಮೂಲ ಪದಕ್ಕೆ ಹಲವು ಪ್ರತ್ಯಯಗಳು ಸೇರಿ ಪದದ ರೂಪ ದೊಡ್ಡದಾಗುತ್ತಾ ಹೋಗುತ್ತದೆ. ಇದನ್ನು ಯಂತ್ರಗಳು ಹೇಗೆ ನಿಭಾಯಿಸುತ್ತವೆ ಎಂಬುದು ಕುತೂಹಲಕಾರಿ:

1. ಪದ ವಿಶ್ಲೇಷಣೆಯ ಸವಾಲು (Morphology): ಕನ್ನಡದಲ್ಲಿ "ಬರೆಯುತ್ತಿರಲಿಲ್ಲವಂತೆ" ಎಂಬ ಒಂದೇ ಪದದಲ್ಲಿ ಇಡೀ ವಾಕ್ಯದ ಅರ್ಥ ಅಡಗಿರುತ್ತದೆ.

ವಿಶ್ಲೇಷಣೆ: ಯಂತ್ರವು ಇದನ್ನು ಬರೆ (ಧಾತು) + ಉತ್ತ (ವರ್ತಮಾನ) + ಇರು (ಸಹಾಯಕ) + ಇಲ್ಲ (ನಕಾರಾತ್ಮಕ) + ಅಂತೆ (ಐತಿಹ್ಯ) ಎಂದು ಬಿಡಿಸಿ ನೋಡಬೇಕು. ಇದಕ್ಕೆ ಕನ್ನಡದಲ್ಲಿ 'Morphological Analyzer' ಎಂಬ ಉಪಕರಣ ಬಳಸಲಾಗುತ್ತದೆ.

2. ವಿಭಕ್ತಿ ಪ್ರತ್ಯಯಗಳ ಗುರುತಿಸುವಿಕೆ: ಕನ್ನಡದಲ್ಲಿ ಕಾರಕಗಳು (Case Markers) ಬಹಳ ಮುಖ್ಯ.

ಉದಾಹರಣೆ: "ರಾಮನು ಹಣ್ಣನ್ನು ತಿಂದನು" ಇಲ್ಲಿ 'ನು' ಮತ್ತು 'ನ್ನು' ಪ್ರತ್ಯಯಗಳನ್ನು ನೋಡಿ ಯಂತ್ರವು ಯಾರು ಯಾರನ್ನು ತಿಂದರು ಎಂದು ನಿರ್ಧರಿಸುತ್ತದೆ. ಇಂಗ್ಲಿಷ್‌ನಲ್ಲಿ 'by' ಅಥವಾ 'to' ಎಂಬ ಪ್ರತ್ಯಯಗಳು ಪದದ ಹೊರಗಿರುತ್ತವೆ, ಆದರೆ ಕನ್ನಡದಲ್ಲಿ ಪದದ ಒಳಗೇ ಇರುವುದರಿಂದ NLP ಗೆ ಇದು ಸವಾಲಿನ ಕೆಲಸ.

ಲಿಪಿ ಮತ್ತು ಫಾಂಟ್ ಪರಿವರ್ತನೆ (Unicode): ಹಿಂದೆ ಕನ್ನಡದಲ್ಲಿ ಟೈಪ್ ಮಾಡಲು ಬೇರೆ ಬೇರೆ ಸಾಫ್ಟ್‌ವೇರ್‌ಗಳ ಅಗತ್ಯವಿತ್ತು (ನುಡಿ, ಬರಹ). ಆದರೆ ಈಗ ಯೂನಿಕೋಡ್ (Unicode) ಬಂದ ನಂತರ, ಜಗತ್ತಿನ ಯಾವುದೇ ಮೂಲೆಯಲ್ಲಿರುವ ಯಂತ್ರಕ್ಕೂ ಕನ್ನಡದ ಅಕ್ಷರ 'ಅ' ಎಂದರೆ ಒಂದು ನಿರ್ದಿಷ್ಟ ಸಂಖ್ಯೆಯಾಗಿ (Code point) ಅರ್ಥವಾಗುತ್ತದೆ. ಇದು ಕನ್ನಡದ ಡಿಜಿಟಲೀಕರಣಕ್ಕೆ ದೊಡ್ಡ ಶಕ್ತಿ ನೀಡಿದೆ.

ಆಪ್ಟಿಕಲ್ ಕ್ಯಾರೆಕ್ಟರ್ ರೆಕಗ್ನಿಷನ್ (OCR): ಹಳೆಯ ಕನ್ನಡ ಪುಸ್ತಕಗಳನ್ನು ಸ್ಕ್ಯಾನ್ ಮಾಡಿ ಅದನ್ನು ಡಿಜಿಟಲ್ ಪಠ್ಯವಾಗಿ (Editable Text) ಬದಲಾಯಿಸುವ ತಂತ್ರಜ್ಞಾನ ಇಂದು ಸುಧಾರಿಸಿದೆ. ಇದು ಹಳೆಯ ಸಾಹಿತ್ಯವನ್ನು ಉಳಿಸಿಕೊಳ್ಳಲು ಸಹಕಾರಿಯಾಗಿದೆ.

5. ಪ್ರಾಯೋಗಿಕ ಅನ್ವಯಗಳು:

- ಕನ್ನಡ ಕೀಬೋರ್ಡ್‌ಗಳು: ನಾವು ಟೈಪ್ ಮಾಡುವಾಗ ಮುಂದಿನ ಪದವನ್ನು ಊಹಿಸುವ (Predictive text) ತಂತ್ರಜ್ಞಾನ.
- ಧ್ವನಿ ಟೈಪಿಂಗ್: ಗೂಗಲ್‌ನಲ್ಲಿ ನಾವು ಕನ್ನಡದಲ್ಲಿ ಮಾತನಾಡಿದರೆ ಅದು ಪಠ್ಯವಾಗಿ ಮೂಡುವುದು.

NLP ತಂತ್ರಜ್ಞಾನವು ಪದಗಳಿಂದ ವಾಕ್ಯಕ್ಕೆ, ವಾಕ್ಯದಿಂದ ಅರ್ಥಕ್ಕೆ ಮತ್ತು ಧ್ವನಿಯಿಂದ ಲಿಪಿಗೆ ಸಾಗುವ ಒಂದು ಹಂತ ಹಂತದ ವೈಜ್ಞಾನಿಕ ಪ್ರಕ್ರಿಯೆ. ಇದು ಯಂತ್ರಗಳಿಗೆ 'ಮಾನವ ಸಹಜ' ಬುದ್ಧಿವಂತಿಕೆಯನ್ನು ನೀಡುವಲ್ಲಿ ಪ್ರಮುಖ ಪಾತ್ರ ವಹಿಸುತ್ತಿದೆ.

ಕೃತಕ ಬುದ್ಧಿಮತ್ತೆಯ ವಿಕಸನ: ಭಾಷಾ ತಂತ್ರಜ್ಞಾನದ ಹೊಸ ಆಯಾಮ

ಭಾಷಾ ತಂತ್ರಜ್ಞಾನವು 'ಕೇವಲ ಶಬ್ದಗಳನ್ನು ಜೋಡಿಸುವ' ಹಂತದಿಂದ 'ಭಾವನೆಗಳನ್ನು ಗ್ರಹಿಸುವ' ಹಂತಕ್ಕೆ ಬೆಳೆದು ಬಂದ ದಾರಿಯನ್ನು ಈ ಕೆಳಗಿನಂತೆ ವಿಶ್ಲೇಷಿಸಬಹುದು:

1. ತಾಂತ್ರಿಕ ವಿಕಸನದ ಮೂರು ಘಟ್ಟಗಳು:

AI ಬೆಳೆದು ಬಂದ ಹಾದಿಯನ್ನು ನಾವು ಮೂರು ಮುಖ್ಯ ಭಾಗಗಳಾಗಿ

ವಿಂಗಡಿಸಬಹುದು:

ಸ್ಥಿರ ನಿಯಮಗಳ ಕಾಲ (Symbolic AI): 1950 ರಿಂದ 1990ರ ದಶಕದವರೆಗೆ. ಇಲ್ಲಿ ತಜ್ಞರು ಭಾಷೆಯ ಪ್ರತಿಯೊಂದು ವ್ಯಾಕರಣ ನಿಯಮವನ್ನು ಕಂಪ್ಯೂಟರ್‌ಗೆ ಬರೆದುಕೊಡಬೇಕಿತ್ತು. ಇದು ತುಂಬಾ ನಿಧಾನಗತಿಯ ಮತ್ತು ಮಿತಿಗಳಿರುವ ವ್ಯವಸ್ಥೆಯಾಗಿತ್ತು.

ಸಂಖ್ಯಾಶಾಸ್ತ್ರದ ಕಾಲ (Statistical NLP): 2000ದ ದಶಕದಲ್ಲಿ ಆರಂಭವಾಯಿತು. ಇಲ್ಲಿ ಯಂತ್ರಗಳು ಗಣಿತದ ಸಂಭವನೀಯತೆ (Probability) ಆಧಾರದ ಮೇಲೆ ಭಾಷಾಂತರಿಸಲು ಕಲಿತವು. ಉದಾಹರಣೆಗೆ, "ನಮಸ್ಕಾರ" ಎಂಬ ಪದದ ನಂತರ "ಹೇಗಿದ್ದೀರಿ" ಬರುವ ಸಾಧ್ಯತೆ ಹೆಚ್ಚು ಎಂದು ಲೆಕ್ಕಾಚಾರ ಹಾಕುವುದು.

ನರಮಂಡಲದ ಕ್ರಾಂತಿ (Neural Revolution): 2010ರ ನಂತರ Deep Learning ಪರಿಚಯವಾಯಿತು. ಇದು ಮಾನವ ಮೆದುಳಿನಂತೆ ಕಾರ್ಯನಿರ್ವಹಿಸುವ 'ಕೃತಕ ನರಮಂಡಲಗಳನ್ನು' (Artificial Neural Networks) ಬಳಸುತ್ತದೆ. ಇದು ಭಾಷೆಯ ಅರ್ಥವನ್ನು ಆಳವಾಗಿ ಗ್ರಹಿಸಲು ನೆರವಾಯಿತು.

2. ಯಂತ್ರ ಕಲಿಕೆ (Machine Learning) – 'ಕಲಿಸುವ' ಬದಲು 'ಕಲಿಯುವ' ಪ್ರಕ್ರಿಯೆ:

ಯಂತ್ರ ಕಲಿಕೆಯ ಪ್ರವೇಶದಿಂದ ನಾವು ಸಾಫ್ಟ್‌ವೇರ್‌ಗಳಿಗೆ "ಹೀಗೆ ಮಾಡು" ಎಂದು ಹೇಳುವ ಅಗತ್ಯವಿಲ್ಲದಂತಾಯಿತು.

ವಿಶ್ಲೇಷಣೆ: ನಾವು ಸಾವಿರಾರು ಕನ್ನಡ ಮತ್ತು ಇಂಗ್ಲಿಷ್ ಲೇಖನಗಳನ್ನು ಯಂತ್ರಕ್ಕೆ ನೀಡಿದಾಗ, ಅದು ತಾನಾಗಿಯೇ ಪದಗಳ ನಡುವಿನ ಸಂಬಂಧವನ್ನು ಕಂಡುಕೊಳ್ಳುತ್ತದೆ.

ವೈಶಿಷ್ಟ್ಯ : ಇದು ಸ್ವಯಂ-ತಿದ್ದಿಕೊಳ್ಳುವ (Self-correcting) ಸಾಮರ್ಥ್ಯ ಹೊಂದಿದೆ. ಬಳಕೆದಾರರು ತಪ್ಪುಗಳನ್ನು ಸರಿಪಡಿಸಿದಂತೆಲ್ಲಾ ಯಂತ್ರದ ನಿಖರತೆ ಹೆಚ್ಚುತ್ತಾ ಹೋಗುತ್ತದೆ.

3. ನ್ಯೂರಲ್ ನೆಟ್‌ವರ್ಕ್ಸ್ ಮತ್ತು ವರ್ಡ್ ಎಂಬೆಡಿಂಗ್ಸ್ (Word Embeddings):

ಭಾಷೆಯನ್ನು ಯಂತ್ರಗಳು ಗಣಿತದ ರೂಪದಲ್ಲಿ ನೋಡುತ್ತವೆ. ನ್ಯೂರಲ್ ನೆಟ್‌ವರ್ಕ್‌ಗಳು ಪ್ರತಿಯೊಂದು ಪದಕ್ಕೂ ಒಂದು ಗಣಿತೀಯ ಸ್ಥಾನವನ್ನು (Vector) ನೀಡುತ್ತವೆ.

ಆಳವಾದ ನೋಟ: ಉದಾಹರಣೆಗೆ, 'ರಾಜ' ಮತ್ತು 'ರಾಣಿ' ಎಂಬ ಪದಗಳ ನಡುವಿನ ಸಂಬಂಧವು 'ಪುರುಷ' ಮತ್ತು 'ಮಹಿಳೆ' ನಡುವಿನ ಸಂಬಂಧಕ್ಕೆ ಸಮನಾಗಿರುತ್ತದೆ ಎಂದು ಯಂತ್ರವು ಗಣಿತದ ಮೂಲಕ ಪತ್ತೆ ಮಾಡುತ್ತದೆ. ಇದನ್ನು Semantic Space ಎನ್ನಲಾಗುತ್ತದೆ. ಇದರಿಂದ ಭಾಷೆಯ ಒಳಗಿನ ಸಂಕೀರ್ಣ

ಸಂಬಂಧಗಳು ಯಂತ್ರಕ್ಕೆ ಅರ್ಥವಾಗುತ್ತವೆ.

4. ಜನರೇಟಿವ್ AI ಮತ್ತು ಟ್ರಾನ್ಸ್‌ಫಾರ್ಮರ್ ಮಾದರಿಗಳು:

ಇಂದಿನ ಚಾಟ್‌ಜಿಪಿಟಿ (ChatGPT) ನಂತಹ ತಂತ್ರಜ್ಞಾನದ ಹಿಂದೆ 'Transformer' ಎಂಬ ಕ್ರಾಂತಿಕಾರಿ ತಂತ್ರಜ್ಞಾನವಿದೆ.

ಗಮನದ ಶಕ್ತಿ (Attention Mechanism): ಈ ಮಾಡೆಲ್‌ಗಳು ಒಂದು ವಾಕ್ಯವನ್ನು ಓದುವಾಗ ಪ್ರತಿಯೊಂದು ಪದಕ್ಕೂ ಎಷ್ಟು ಮಹತ್ವ ಕೊಡಬೇಕು ಎಂಬುದನ್ನು ಅರಿಯುತ್ತವೆ. ಉದಾಹರಣೆಗೆ: "ನದಿಯ ದಡದಲ್ಲಿದ್ದ ಆ ಮರದ ಹಣ್ಣು ಕೆಂಪಗಿತ್ತು" – ಇಲ್ಲಿ 'ಕೆಂಪಗಿತ್ತು' ಎಂಬುದು ಮರಕ್ಕೋ ಅಥವಾ ಹಣ್ಣಿಗೋ ಎಂಬ ಗೊಂದಲವನ್ನು ಈ ತಂತ್ರಜ್ಞಾನ ನಿವಾರಿಸುತ್ತದೆ.

ಸೃಜನಶೀಲತೆ: ಇದು ಕೇವಲ ಇರುವ ಮಾಹಿತಿಯನ್ನು ನೀಡುವುದಿಲ್ಲ, ಬದಲಿಗೆ ಹೊಸದಾಗಿ ಕವಿತೆ, ಕಥೆ ಅಥವಾ ಇಮೇಲ್‌ಗಳನ್ನು ಸೃಷ್ಟಿಸುತ್ತದೆ.

5. ಈ ವಿಕಸನದ ಸಾಮಾಜಿಕ ಪ್ರಭಾವ (Analytical Impact):

ಭಾಷಾ ಅಡೆತಡೆಗಳ ನಿವಾರಣೆ: ಪ್ರಾದೇಶಿಕ ಭಾಷೆಗಳಿಗೂ ಇಂದು ಉತ್ತಮ ತಂತ್ರಜ್ಞಾನ ಲಭ್ಯವಿರುವುದರಿಂದ, ಜ್ಞಾನವು ಕೇವಲ ಇಂಗ್ಲಿಷ್ ತಿಳಿದವರಿಗೆ ಸೀಮಿತವಾಗಿಲ್ಲ.

ವೈಯಕ್ತಿಕ ಸಹಾಯಕರು: ಸಿರಿ (Siri) ಅಥವಾ ಅಲೆಕ್ಸಾದಂತಹ ಸಾಧನಗಳು ನಮ್ಮ ಭಾಷೆಯ ಶೈಲಿಯನ್ನು ಕಲಿತು ನಮಗೆ ತಕ್ಕಂತೆ ಪ್ರತಿಕ್ರಿಯಿಸುತ್ತಿವೆ.

AI ನ ವಿಕಸನವು ಭಾಷೆಯನ್ನು ಯಂತ್ರಕ್ಕೆ ಕೇವಲ ಒಂದು ಡೇಟಾ ಆಗಿ ನೋಡದೆ, ಮಾನವನ ಆಲೋಚನೆಗಳನ್ನು ಅಭಿವ್ಯಕ್ತಿಸುವ ಒಂದು ಮಾಧ್ಯಮವಾಗಿ ನೋಡುವಂತೆ ಮಾಡಿದೆ. ಇದು "ಮಾಹಿತಿ ಯುಗ" ದಿಂದ "ಅರಿವಿನ ಯುಗ"ಕ್ಕೆ (Era of Understanding) ನಾವು ದಾಪುಗಾಲು ಹಾಕಲು ನೆರವಾಗಿದೆ.

ಆಧುನಿಕ ಯುಗದಲ್ಲಿ ಭಾಷಾ ತಂತ್ರಜ್ಞಾನ ಮತ್ತು ಅದರ ವ್ಯಾಪ್ತಿ

ಖಂಡಿತ, ಭಾಷಾ ತಂತ್ರಜ್ಞಾನವು (Language Technology) ಕೇವಲ ಆಪ್‌ಗಳಿಗೆ ಸೀಮಿತವಾಗದೆ, ಇಂದಿನ ಸಮಾಜದ ಆರ್ಥಿಕ, ಸಾಮಾಜಿಕ ಮತ್ತು ಶೈಕ್ಷಣಿಕ ವ್ಯವಸ್ಥೆಯ ಅವಿಭಾಜ್ಯ ಅಂಗವಾಗಿದೆ. ನೀವು ಉಲ್ಲೇಖಿಸಿದ ಅಂಶಗಳನ್ನು ಆಧಾರವಾಗಿಟ್ಟುಕೊಂಡು, ಈ ತಂತ್ರಜ್ಞಾನದ ಪ್ರಭಾವವನ್ನು ಇನ್ನೂ ವಿಸ್ತಾರವಾಗಿ ಮತ್ತು ವಿಶ್ಲೇಷಣಾತ್ಮಕವಾಗಿ ಈ ಕೆಳಗಿನಂತೆ ವಿವರಿಸಬಹುದು:

ಭಾಷೆಯು ಮಾನವಕುಲದ ಸಂವಹನದ ಜೀವಾಳ. ತಂತ್ರಜ್ಞಾನ ಬೆಳೆದಂತೆ ಮನುಷ್ಯ ಮತ್ತು ಯಂತ್ರಗಳ ನಡುವಿನ ಸಂಭಾಷಣೆಯನ್ನು ಸುಲಭಗೊಳಿಸಲು ಹುಟ್ಟಿಕೊಂಡಿದ್ದೇ 'ಭಾಷಾ ತಂತ್ರಜ್ಞಾನ'. ಇದು ಕಂಪ್ಯೂಟರ್ ವಿಜ್ಞಾನ, ಭಾಷಾಶಾಸ್ತ್ರ

ಮತ್ತು ಕೃತಕ ಬುದ್ಧಿಮತ್ತೆಯ (AI) ಸಮ್ಮಿಲನವಾಗಿದೆ. ಇಂದು ನಮ್ಮ ಕೈಯಲ್ಲಿರುವ ಸ್ಮಾರ್ಟ್‌ಫೋನ್‌ನಿಂದ ಹಿಡಿದು ದೊಡ್ಡ ಸಂಶೋಧನಾ ಕೇಂದ್ರಗಳವರೆಗೆ ಈ ತಂತ್ರಜ್ಞಾನವು ಮನುಷ್ಯನ 'ಅದೃಶ್ಯ ಶಕ್ತಿ'ಯಾಗಿ ಕಾರ್ಯನಿರ್ವಹಿಸುತ್ತಿದೆ.

ಶೈಕ್ಷಣಿಕ ಮತ್ತು ಬೌದ್ಧಿಕ ಕ್ರಾಂತಿ: ಶಿಕ್ಷಣ ಕ್ಷೇತ್ರದಲ್ಲಿ ಭಾಷಾ ತಂತ್ರಜ್ಞಾನವು ಕ್ರಾಂತಿಕಾರಿ ಬದಲಾವಣೆ ತಂದಿದೆ. ಇದು ಜ್ಞಾನವನ್ನು ಕೇವಲ ಇಂಗ್ಲಿಷ್ ಭಾಷೆಗೆ ಸೀಮಿತಗೊಳಿಸದೆ ಪ್ರಾದೇಶಿಕ ಭಾಷೆಗಳಿಗೂ ವಿಸ್ತರಿಸಿದೆ. AI ಆಧಾರಿತ ಭಾಷಾಂತರ ಸಾಧನಗಳು ಸಂಕೀರ್ಣ ವೈಜ್ಞಾನಿಕ ವಿಷಯಗಳನ್ನು ಕ್ಷಣಾರ್ಧದಲ್ಲಿ ಕನ್ನಡಕ್ಕೆ ತರ್ಜುಮೆ ಮಾಡುತ್ತಿವೆ. ಅಲ್ಲದೆ, 'ಡ್ಯುಯೋಲಿಂಗೋ' ನಂತಹ ಆಪ್‌ಗಳು ವೈಯಕ್ತಿಕ ಶಿಕ್ಷಕರಂತೆ ಕಾರ್ಯನಿರ್ವಹಿಸಿ, ಉಚ್ಚಾರಣೆ ಮತ್ತು ವ್ಯಾಕರಣವನ್ನು ತಿದ್ದುವ ಮೂಲಕ ಭಾಷಾ ಕಲಿಕೆಯನ್ನು ಸುಲಭಗೊಳಿಸಿವೆ. ಇದರಿಂದ 'ಜ್ಞಾನದ ಪ್ರಜಾಪ್ರಭುತ್ವೀಕರಣ' ಸಾಧ್ಯವಾಗಿದೆ.

ಆರೋಗ್ಯ ಮತ್ತು ಸಾಮಾಜಿಕ ಹಿತರಕ್ಷಣೆ: ಆರೋಗ್ಯ ಕ್ಷೇತ್ರದಲ್ಲಿ ಈ ತಂತ್ರಜ್ಞಾನವು ಜೀವ ಉಳಿಸುವಲ್ಲಿ ಸಹಕಾರಿಯಾಗಿದೆ. ವೈದ್ಯಕೀಯ ವರದಿಗಳ ವಿಶ್ಲೇಷಣೆ ಮತ್ತು ವೈದ್ಯರ ಮೌಖಿಕ ಸಲಹೆಗಳನ್ನು ತಕ್ಷಣ ಪಠ್ಯ ರೂಪಕ್ಕೆ ತರುವ (Medical Transcription) ಪ್ರಕ್ರಿಯೆಯಲ್ಲಿ ಇದು ಬಳಕೆಯಾಗುತ್ತಿದೆ. ವಿಶೇಷವಾಗಿ ಮಾನಸಿಕ ಆರೋಗ್ಯದ ಕ್ಷೇತ್ರದಲ್ಲಿ, ಒಬ್ಬ ವ್ಯಕ್ತಿಯ ಸಂವಹನ ಶೈಲಿಯನ್ನು ವಿಶ್ಲೇಷಿಸಿ (Sentiment Analysis), ಅವರು ಖಿನ್ನತೆಯಲ್ಲಿದ್ದಾರೆಯೇ ಎಂಬುದನ್ನು ಪತ್ತೆಹಚ್ಚುವಲ್ಲಿ AI ಮಹತ್ತರ ಪಾತ್ರ ವಹಿಸುತ್ತಿದೆ.

ಆಡಳಿತ ಮತ್ತು ವ್ಯವಹಾರದಲ್ಲಿನ ಬಳಕೆ: ಸರ್ಕಾರ ಮತ್ತು ನ್ಯಾಯಾಂಗ ವ್ಯವಸ್ಥೆಯು ಜನಸಾಮಾನ್ಯರಿಗೆ ಹತ್ತಿರವಾಗಲು ಭಾಷಾ ತಂತ್ರಜ್ಞಾನವನ್ನು ಬಳಸುತ್ತಿವೆ. ಉದಾಹರಣೆಗೆ, ನ್ಯಾಯಾಲಯದ ತೀರ್ಪುಗಳನ್ನು ಪ್ರಾದೇಶಿಕ ಭಾಷೆಗಳಿಗೆ ಭಾಷಾಂತರಿಸುವ 'ಸುವಾಸ್' (SUVAS) ತಂತ್ರಜ್ಞಾನವು ಕಾನೂನು ಅರಿವು ಮೂಡಿಸಲು ನೆರವಾಗುತ್ತಿದೆ. ವ್ಯಾಪಾರ ಕ್ಷೇತ್ರದಲ್ಲಿ, ಗ್ರಾಹಕರ ಪ್ರಶ್ನೆಗಳಿಗೆ ತಕ್ಷಣ ಉತ್ತರಿಸುವ 'ಚಾಟ್‌ಬಾಟ್‌ಗಳು' ಮತ್ತು ಗ್ರಾಹಕರ ಅಗತ್ಯಕ್ಕೆ ತಕ್ಕಂತೆ ಉತ್ಪನ್ನಗಳನ್ನು ಶಿಫಾರಸು ಮಾಡುವ ವ್ಯವಸ್ಥೆಗಳು ಇ-ಕಾಮರ್ಸ್‌ನ ಚಿತ್ರಣವನ್ನೇ ಬದಲಿಸಿವೆ.

ಮಾಧ್ಯಮ ಮತ್ತು ಮನರಂಜನೆಯ ಹೊಸ ಆಯಾಮ: ನಾವಿಂದು ಯೂಟ್ಯೂಬ್ ಅಥವಾ ಓಟಿಟಿ ವೇದಿಕೆಗಳಲ್ಲಿ ನೋಡುವ ವಿಡಿಯೋಗಳಿಗೆ ಸ್ವಯಂಚಾಲಿತ ಉಪಶೀರ್ಷಿಕೆಗಳು (Auto-Subtitles) ಈ ತಂತ್ರಜ್ಞಾನದ ಕೊಡುಗೆ. ಸಾವಿರಾರು ಪುಟಗಳ ಸುದ್ದಿಯ ಸಾರಾಂಶವನ್ನು ಕೆಲವೇ ಸಾಲುಗಳಲ್ಲಿ ನೀಡುವ ಶಕ್ತಿಯನ್ನು AI ಹೊಂದಿದೆ. ಸಾಹಿತ್ಯ ಕ್ಷೇತ್ರದಲ್ಲಿ 'ಜನರೇಟಿವ್ AI' ಕವಿತೆ ಮತ್ತು ಲೇಖನಗಳನ್ನು ಬರೆಯುವ ಮೂಲಕ ಸೃಜನಶೀಲತೆಗೆ ಹೊಸ ಆಯಾಮ ನೀಡುತ್ತಿದೆ.

ಸವಾಲುಗಳು ಮತ್ತು ನೈತಿಕತೆ: ಯಾವುದೇ ತಂತ್ರಜ್ಞಾನದಂತೆ ಇದರಲ್ಲೂ

ಕೆಲವು ಸವಾಲುಗಳಿವೆ. AI ಮಾದರಿಗಳು ದತ್ತಾಂಶದಲ್ಲಿರುವ ಪೂರ್ವಾಗ್ರಹಗಳನ್ನು (Bias) ಕಲಿಯುವ ಸಾಧ್ಯತೆ ಇರುತ್ತದೆ. ಅಲ್ಲದೆ, ಸದಾ ನಮ್ಮ ಧ್ವನಿಯನ್ನು ಆಲಿಸುವ ಸಾಧನಗಳು ಖಾಸಗಿತನದ (Privacy) ಹರಣಕ್ಕೆ ದಾರಿ ಮಾಡಿಕೊಡಬಹುದು ಎಂಬ ಆತಂಕವೂ ಇದೆ. ಈ ತಂತ್ರಜ್ಞಾನವನ್ನು ಬಳಸುವಾಗ ನೈತಿಕ ಎಚ್ಚರಿಕೆ ವಹಿಸುವುದು ಅತ್ಯಗತ್ಯ.

ಒಟ್ಟಾರೆಯಾಗಿ ಹೇಳುವುದಾದರೆ, ಭಾಷಾ ತಂತ್ರಜ್ಞಾನವು ಭಾಷಾ ಅಡೆತಡೆಗಳನ್ನು ಅಳಿಸಿ ಜಗತ್ತನ್ನು ಒಂದುಗೂಡಿಸುತ್ತಿದೆ. ಇದು ಕೇವಲ ಒಂದು ಸೌಲಭ್ಯವಲ್ಲ, ಬದಲಿಗೆ ಜ್ಞಾನ ಮತ್ತು ಸಂವಹನದ ಸೇತುವೆಯಾಗಿದೆ. ಇದರ ಸವಾಲುಗಳನ್ನು ಮೀರಿ, ಮಾನವೀಯ ಮೌಲ್ಯಗಳೊಂದಿಗೆ ಈ ತಂತ್ರಜ್ಞಾನವನ್ನು ಅಳವಡಿಸಿಕೊಂಡರೆ ಮುಂದಿನ ದಿನಗಳಲ್ಲಿ ಮಾಹಿತಿ ಮತ್ತು ಸಂವಹನ ಕ್ಷೇತ್ರವು ಇನ್ನಷ್ಟು ಉತ್ತುಂಗಕ್ಕೇರುವುದರಲ್ಲಿ ಸಂಶಯವಿಲ್ಲ.

ಪ್ರಾದೇಶಿಕ ಭಾಷೆಗಳು ಮತ್ತು ತಾಂತ್ರಿಕ ಸವಾಲುಗಳು:

ಜಾಗತಿಕ ಮಟ್ಟದಲ್ಲಿ ಕೃತಕ ಬುದ್ಧಿಮತ್ತೆ (AI) ಮತ್ತು ಭಾಷಾ ತಂತ್ರಜ್ಞಾನವು ವೇಗವಾಗಿ ಬೆಳೆಯುತ್ತಿದ್ದರೂ, ಪ್ರಾದೇಶಿಕ ಭಾಷೆಗಳ ವಿಷಯಕ್ಕೆ ಬಂದಾಗ ಅದರ ಹಾದಿ ಸುಗಮವಾಗಿಲ್ಲ. ಕನ್ನಡದಂತಹ ಸಾವಿರಾರು ವರ್ಷಗಳ ಇತಿಹಾಸವಿರುವ ಭಾಷೆಗಳಿಗೆ ಆಧುನಿಕ ತಂತ್ರಜ್ಞಾನವನ್ನು ಅಳವಡಿಸುವುದು ಕೇವಲ ತಾಂತ್ರಿಕ ಪ್ರಕ್ರಿಯೆಯಲ್ಲ, ಅದೊಂದು ಸಾಂಸ್ಕೃತಿಕ ಸವಾಲು ಕೂಡ ಹೌದು. ಈ ಹಾದಿಯಲ್ಲಿರುವ ಪ್ರಮುಖ ಅಡೆತಡೆಗಳನ್ನು ಈ ಕೆಳಗಿನಂತೆ ವಿಶ್ಲೇಷಿಸಬಹುದು.

ದತ್ತಾಂಶದ ಕೊರತೆ: AI ಮಾದರಿಗಳ ಯಶಸ್ಸು ಅವುಗಳಿಗೆ ನೀಡುವ ದತ್ತಾಂಶದ (Data) ಮೇಲೆ ಅವಲಂಬಿತವಾಗಿದೆ. ಇಂಗ್ಲಿಷ್ ಭಾಷೆಯಲ್ಲಿ ಕೋಟ್ಯಂತರ ಡಿಜಿಟಲ್ ಪುಸ್ತಕಗಳು, ಲೇಖನಗಳು ಮತ್ತು ಸಂಭಾಷಣೆಗಳು ಲಭ್ಯವಿವೆ. ಆದರೆ, ಕನ್ನಡದಲ್ಲಿ ಡಿಜಿಟಲ್ ರೂಪದ ಬರಹಗಳು ತುಲನಾತ್ಮಕವಾಗಿ ಕಡಿಮೆ.

ಯಂತ್ರ ಕಲಿಕೆಗೆ (Machine Learning) ಬೇಕಾದ 'ಗುಣಮಟ್ಟದ ದತ್ತಾಂಶ' ಸಿಗುವುದು ಕಷ್ಟವಾಗಿದೆ. ಸಾಮಾಜಿಕ ಜಾಲತಾಣಗಳಲ್ಲಿನ ಕನ್ನಡವು ಹೆಚ್ಚಾಗಿ 'ತಪ್ಪು ವ್ಯಾಕರಣ' ಅಥವಾ 'ಇಂಗ್ಲಿಷ್ ಬೆರೆತ ಕನ್ನಡ' (Kanglish) ಆಗಿರುವುದರಿಂದ, ಯಂತ್ರಗಳಿಗೆ ಶುದ್ಧ ಕನ್ನಡವನ್ನು ಕಲಿಸಲು ಬೇಕಾದ ಗ್ರಂಥಾಲಯಗಳ ಡಿಜಿಟಲೀಕರಣ ಇನ್ನೂ ವೇಗ ಪಡೆಯಬೇಕಿದೆ.

ವ್ಯಾಕರಣ ಮತ್ತು ಭಾಷಾ ರಚನೆಯ ಸಂಕೀರ್ಣತೆ: ಕನ್ನಡವು ಒಂದು 'ಅಂಟಿಸುವ ಭಾಷೆ' (Agglutinative Language). ಇಲ್ಲಿ ಒಂದು ಸಣ್ಣ ಪ್ರತ್ಯಯವು ಇಡೀ ವಾಕ್ಯದ ಅರ್ಥವನ್ನೇ ಬದಲಿಸಬಲ್ಲದು.

ಇಂಗ್ಲಿಷ್‌ನಲ್ಲಿ ಪದಗಳು ಬಿಡಿಬಿಡಿಯಾಗಿರುತ್ತವೆ (ಉದಾ: 'To the house'). ಆದರೆ ಕನ್ನಡದಲ್ಲಿ ಅದು 'ಮನೆಗೆ' ಎಂದಾಗುತ್ತದೆ. ಸಂಧಿಗಾರ್ಯಗಳು ನಡೆದಾಗ

ಪದಗಳ ರೂಪವೇ ಬದಲಾಗುತ್ತದೆ (ಉದಾ: ಮಳೆ + ಕಾಲ = ಮಳೆಗಾಲ). ಇಂತಹ ವ್ಯಾಕರಣದ ವೈವಿಧ್ಯತೆಯನ್ನು ಮತ್ತು ಸಮಾಸಗಳ ಸಂಕೀರ್ಣತೆಯನ್ನು ಡಿಜಿಟಲ್ ರೂಪಕ್ಕೆ ತರುವುದು AI ಸಂಶೋಧಕರಿಗೆ ದೊಡ್ಡ ಸವಾಲಾಗಿದೆ.

ಪ್ರಾದೇಶಿಕ ಉಪಭಾಷೆಗಳು ಮತ್ತು ವೈವಿಧ್ಯತೆ: "ಹನ್ನೆರಡು ಗಾವುದ ಹೋದರೆ ಭಾಷೆ ಬದಲಾಗುತ್ತದೆ" ಎಂಬ ಗಾದೆಯಂತೆ ಕನ್ನಡದಲ್ಲಿ ಪ್ರಾದೇಶಿಕ ವೈವಿಧ್ಯತೆ ಅಪಾರ.

ಮಂಗಳೂರು ಕನ್ನಡ, ಧಾರವಾಡ ಕನ್ನಡ, ಮೈಸೂರು ಕನ್ನಡ ಮತ್ತು ಗುಲ್ಬರ್ಗಾ ಕನ್ನಡದ ಲಹರಿಗಳು ವಿಭಿನ್ನವಾಗಿವೆ. ಇದರ ಜೊತೆಗೆ ಹಳೇಗನ್ನಡ ಮತ್ತು ನಡುಗನ್ನಡದ ಸಾಹಿತ್ಯಿಕ ಭಾಷೆ ಬೇರೆಯೇ ಇದೆ. ಒಂದು ಧ್ವನಿ ಆಧಾರಿತ AI (Voice AI) ಇವೆಲ್ಲವನ್ನೂ ಒಂದೇ ಸಮನಾಗಿ ಗ್ರಹಿಸುವುದು ಕಷ್ಟ. ಯಾವುದೇ ಒಂದು ಉಪಭಾಷೆಗೆ ಹೆಚ್ಚಿನ ಆದ್ಯತೆ ನೀಡಿದರೆ, ಉಳಿದ ಶೈಲಿಗಳು ತಾಂತ್ರಿಕವಾಗಿ ಹಿಂದೆ ಉಳಿಯುವ ಅಪಾಯವಿರುತ್ತದೆ.

ಲಿಪಿ, ಟೈಪಿಂಗ್ ಮತ್ತು ಪ್ರಮಾಣೀಕರಣ (Standardization): ಅನೇಕ ವರ್ಷಗಳ ಕಾಲ ಕನ್ನಡದಲ್ಲಿ ವಿವಿಧ ಫಾಂಟ್‌ಗಳು ಮತ್ತು ಕೀಬೋರ್ಡ್ ಲೇಔಟ್‌ಗಳು (ನುಡಿ, ಬರಹ, ಯೂನಿಕೋಡ್) ಬಳಕೆಯಲ್ಲಿದ್ದವು.

ಇಂದಿಗೂ ಅನೇಕ ಸರ್ಕಾರಿ ದಾಖಲೆಗಳು ಅಥವಾ ಹಳೆಯ ಕಡತಗಳು ಯೂನಿಕೋಡ್ ಅಲ್ಲದ (Non-Unicode) ಫಾಂಟ್‌ಗಳಲ್ಲಿವೆ. ಇವುಗಳನ್ನು ಯಂತ್ರಗಳು ಓದಲು ಸಾಧ್ಯವಾಗುವುದಿಲ್ಲ. ಸಾಮಾನ್ಯ ಜನರಿಗೆ ಇಂದಿಗೂ ಮೊಬೈಲ್‌ನಲ್ಲಿ ಕನ್ನಡ ಟೈಪಿಂಗ್ ಮಾಡುವುದು ಒಂದು ಹೊರೆಯಾಗಿ ಕಂಡುಬರುತ್ತದೆ, ಇದು ಡಿಜಿಟಲ್ ಲೋಕದಲ್ಲಿ ಕನ್ನಡದ ಬಳಕೆಯನ್ನು ಕಡಿಮೆ ಮಾಡುತ್ತದೆ.

ಆರ್ಥಿಕ ಮತ್ತು ಮಾರುಕಟ್ಟೆಯ ಅಸಮಾನತೆ: ತಂತ್ರಜ್ಞಾನದ ಹೂಡಿಕೆ ಹೆಚ್ಚಾಗಿ ಇಂಗ್ಲಿಷ್ ಅಥವಾ ಮಾರುಕಟ್ಟೆ ದೊಡ್ಡದಾಗಿರುವ ಭಾಷೆಗಳತ್ತ ಹರಿಯುತ್ತದೆ.

ಜಾಗತಿಕ ಕಂಪನಿಗಳಿಗೆ ಕನ್ನಡದಂತಹ ಪ್ರಾದೇಶಿಕ ಭಾಷೆಗಳಿಗೆ ತಂತ್ರಜ್ಞಾನ ರೂಪಿಸುವುದು ಲಾಭದಾಯಕವಾಗಿ ಕಾಣದೇ ಇರಬಹುದು. ಇದರಿಂದಾಗಿ ಅತ್ಯಾಧುನಿಕ AI ಸೌಲಭ್ಯಗಳು ಮೊದಲು ಇಂಗ್ಲಿಷ್‌ನಲ್ಲಿ ಬಂದು, ಹಲವು ವರ್ಷಗಳ ನಂತರ ಕನ್ನಡಕ್ಕೆ ತಲುಪುತ್ತವೆ. ಇದು 'ಡಿಜಿಟಲ್ ಕಂದಕ'ವನ್ನು (Digital Divide) ಸೃಷ್ಟಿಸುತ್ತದೆ.

ಪ್ರಾದೇಶಿಕ ಭಾಷೆಗಳಿಗೆ ಇರುವ ಸವಾಲುಗಳು ಕೇವಲ ತಾಂತ್ರಿಕವಲ್ಲ, ಅವು ಭಾಷೆಯ ಉಳಿವು ಮತ್ತು ಬೆಳವಣಿಗೆಯ ಪ್ರಶ್ನೆಯಾಗಿವೆ. ಈ ಸವಾಲುಗಳನ್ನು ಮೆಟ್ಟಿ ನಿಲ್ಲಲು ಸರ್ಕಾರ, ತಂತ್ರಜ್ಞರು ಮತ್ತು ಸಾಹಿತಿಗಳು ಒಟ್ಟಾಗಿ ಕೆಲಸ ಮಾಡಬೇಕಿದೆ. ಕನ್ನಡದ ದತ್ತಾಂಶವನ್ನು ಹೆಚ್ಚಿಸುವುದು, ತಾಂತ್ರಿಕ ಕನ್ನಡವನ್ನು ಪ್ರಮಾಣೀಕರಿಸುವುದು ಮತ್ತು ಜನಸಾಮಾನ್ಯರಲ್ಲಿ ಡಿಜಿಟಲ್ ಅರಿವು ಮೂಡಿಸುವುದರಿಂದ ಮಾತ್ರ ಕನ್ನಡವು 'ತಾಂತ್ರಿಕ ಭಾಷೆ'ಯಾಗಿ ಜಗತ್ತಿನಲ್ಲಿ ಮಿಂಚಲು ಸಾಧ್ಯ.

ಪ್ರಾದೇಶಿಕ ಭಾಷಾ ತಂತ್ರಜ್ಞಾನದ ಸವಾಲುಗಳಿಗೆ ಪರಿಹಾರ ಮತ್ತು ಭವಿಷ್ಯದ ಹಾದಿ

ಪ್ರಾದೇಶಿಕ ಭಾಷೆಗಳು ಎದುರಿಸುತ್ತಿರುವ ತಾಂತ್ರಿಕ ಸವಾಲುಗಳನ್ನು ಮೆಟ್ಟಿ ನಿಲ್ಲಲು ಕೇವಲ ತಂತ್ರಜ್ಞಾನವಷ್ಟೇ ಸಾಲದು, ಅಲ್ಲಿ ಭಾಷಾ ತಜ್ಞರು ಮತ್ತು ಸಾಮಾನ್ಯ ಜನರ ಸಹಭಾಗಿತ್ವವೂ ಅಗತ್ಯ. ಈ ನಿಟ್ಟಿನಲ್ಲಿ ಅನುಸರಿಸಬಹುದಾದ ಪ್ರಮುಖ ಪರಿಹಾರಗಳು ಇಲ್ಲಿವೆ:

1. ದತ್ತಾಂಶ ಸಂಗ್ರಹಣೆ ಮತ್ತು 'ಕ್ರೌಡ್ ಸೋರ್ಸಿಂಗ್':

ದತ್ತಾಂಶದ ಕೊರತೆಯನ್ನು ನೀಗಿಸಲು ಸರ್ಕಾರ ಮತ್ತು ಸಂಘ ಸಂಸ್ಥೆಗಳು ಮುಂಚೂಣಿಗೆ ಬರಬೇಕಿದೆ.

'ಭಾಷಿಣಿ' (Bhashini) ನಂತಹ ಯೋಜನೆಗಳ ಮೂಲಕ ಸಾಮಾನ್ಯ ಜನರು ತಮ್ಮ ಧ್ವನಿ ಮತ್ತು ಬರಹಗಳನ್ನು ದಾನ ಮಾಡುವಂತೆ ಪ್ರೇರೇಪಿಸಬೇಕು. ಜ್ಞಾನದ ಬೃಹತ್ ಭಂಡಾರವಾದ 'ವಿಕಿಪೀಡಿಯಾ'ದಲ್ಲಿ ಕನ್ನಡದ ಲೇಖನಗಳನ್ನು ಹೆಚ್ಚಿಸುವುದರಿಂದ AI ಮಾದರಿಗಳಿಗೆ ಉತ್ತಮ ತರಬೇತಿ ನೀಡಲು ಸಾಧ್ಯವಾಗುತ್ತದೆ.

2. ಪ್ರಮಾಣಿತ ಮತ್ತು ಮುಕ್ತ ತಂತ್ರಜ್ಞಾನ:

ಕನ್ನಡದ ಟೈಪಿಂಗ್ ಮತ್ತು ಫಾಂಟ್‌ಗಳಲ್ಲಿನ ಗೊಂದಲವನ್ನು ಹೋಗಲಾಡಿಸಲು ಪ್ರಮಾಣೀಕರಣ ಅಗತ್ಯ.

ಎಲ್ಲಾ ಸರ್ಕಾರಿ ಮತ್ತು ಶೈಕ್ಷಣಿಕ ದಾಖಲೆಗಳನ್ನು ಕಡ್ಡಾಯವಾಗಿ 'ಯೂನಿಕೋಡ್' (Unicode) ಗೆ ಪರಿವರ್ತಿಸಬೇಕು. ಮುಕ್ತ ಮೂಲ (Open Source) ತಂತ್ರಜ್ಞಾನಗಳನ್ನು ಉತ್ತೇಜಿಸುವುದರಿಂದ ಸಣ್ಣ ಸಂಶೋಧಕರೂ ಸಹ ಕನ್ನಡದ ಮೇಲೆ ಹೊಸ ಪ್ರಯೋಗಗಳನ್ನು ಮಾಡಲು ಅನುಕೂಲವಾಗುತ್ತದೆ.

3. ಅಡ್ಡ-ಭಾಷಾ ಮಾದರಿಗಳ ಬಳಕೆ:

ಇತ್ತೀಚಿನ AI ತಂತ್ರಜ್ಞಾನದಲ್ಲಿ 'ಜೀರೋ-ಶಾಟ್ ಲರ್ನಿಂಗ್' ಎಂಬ ಮಾದರಿಯಿದೆ.

ಇಲ್ಲಿ ಯಂತ್ರವು ಇಂಗ್ಲಿಷ್‌ನಲ್ಲಿ ಕಲಿತ ಜ್ಞಾನವನ್ನು ಕನ್ನಡಕ್ಕೆ ಅನ್ವಯಿಸಿಕೊಳ್ಳಲು ಪ್ರಯತ್ನಿಸುತ್ತದೆ. ಉದಾಹರಣೆಗೆ, ಇಂಗ್ಲಿಷ್‌ನಲ್ಲಿ 'ವ್ಯಾಕರಣ'ದ ರಚನೆಯನ್ನು ಅರ್ಥಮಾಡಿಕೊಂಡ AI, ಅದನ್ನು ಕನ್ನಡದ ಪದಗಳಿಗೆ ಅನ್ವಯಿಸಿ ಸರಿಯಾದ ವಾಕ್ಯ ರಚಿಸಲು ನೆರವಾಗುತ್ತದೆ. ಇದು ದತ್ತಾಂಶದ ಕೊರತೆ ಇರುವ ಭಾಷೆಗಳಿಗೆ ವರದಾನವಾಗಿದೆ.

4. ಶಿಕ್ಷಣ ಮತ್ತು ಡಿಜಿಟಲ್ ಅರಿವು:

ತಂತ್ರಜ್ಞಾನವು ಎಲ್ಲರಿಗೂ ತಲುಪಬೇಕಾದರೆ, ಕನ್ನಡ ಮಾಧ್ಯಮದ ವಿದ್ಯಾರ್ಥಿಗಳಿಗೆ ಮತ್ತು ಗ್ರಾಮೀಣ ಜನರಿಗೆ ಡಿಜಿಟಲ್ ಕೌಶಲ್ಯಗಳನ್ನು ಕಲಿಸಬೇಕು.

ಕೀಬೋರ್ಡ್ ಬಳಕೆಯ ಬದಲು 'ಧ್ವನಿ ಆಧಾರಿತ ಟೈಪಿಂಗ್' (Voice

Typing) ಬಗ್ಗೆ ಅರಿವು ಮೂಡಿಸಿದರೆ, ಭಾಷೆಯ ಡಿಜಿಟಲ್ ಬಳಕೆ ಹೆಚ್ಚಾಗುತ್ತದೆ. ಇದರಿಂದ ಯಂತ್ರಗಳಿಗೆ ವೈವಿಧ್ಯಮಯವಾದ ಶಬ್ದಗಳ ಮತ್ತು ಉಚ್ಚಾರಣೆಗಳ ದತ್ತಾಂಶ ದೊರೆಯುತ್ತದೆ.

5. ಸರ್ಕಾರ ಮತ್ತು ಖಾಸಗಿ ವಲಯದ ಸಹಯೋಗ:

ಜಾಗತಿಕ ತಂತ್ರಜ್ಞಾನ ಕಂಪನಿಗಳು ಪ್ರಾದೇಶಿಕ ಭಾಷೆಗಳ ಮೇಲೆ ಹೂಡಿಕೆ ಮಾಡುವಂತೆ ಉತ್ತೇಜಿಸಲು ಸರ್ಕಾರವು ಸೂಕ್ತ ನೀತಿಗಳನ್ನು ರೂಪಿಸಬೇಕು. ಕನ್ನಡದ ಸ್ಟಾರ್ಟ್‌ಅಪ್‌ಗಳಿಗೆ ಧನಸಹಾಯ ನೀಡುವುದರಿಂದ ಸ್ಥಳೀಯ ಸಮಸ್ಯೆಗಳಿಗೆ ಸ್ಥಳೀಯವಾಗಿ ತಾಂತ್ರಿಕ ಪರಿಹಾರಗಳನ್ನು ಕಂಡುಕೊಳ್ಳಲು ಸಾಧ್ಯವಾಗುತ್ತದೆ.

ಭವಿಷ್ಯದ ದೃಷ್ಟಿಕೋನ: ಬದಲಾಗಲಿರುವ ಸಂವಹನ ಲೋಕ

ಮುಂದಿನ ದಶಕವು ಭಾಷಾ ತಂತ್ರಜ್ಞಾನದ ಇತಿಹಾಸದಲ್ಲಿ ಒಂದು ಸುವರ್ಣ ಯುಗವಾಗಲಿದೆ. ಇದು ಕೇವಲ ಒಂದು ಭಾಷೆಯಿಂದ ಮತ್ತೊಂದು ಭಾಷೆಗೆ ಪದಗಳನ್ನು ಬದಲಿಸುವ ಹಂತವನ್ನು ಮೀರಿ 'ಸಂವೇದನಾಶೀಲ' ತಂತ್ರಜ್ಞಾನವಾಗಿ ಹೊರಹೊಮ್ಮಲಿದೆ.

ಮಲ್ಟಿಮೋಡಲ್ ಕ್ರಾಂತಿ: ಭವಿಷ್ಯದ AI ಕೇವಲ ಪಠ್ಯ ಅಥವಾ ಧ್ವನಿಗೆ ಸೀಮಿತವಾಗುವುದಿಲ್ಲ. ಅದು ಪಠ್ಯ, ಧ್ವನಿ, ಚಿತ್ರ, ವಿಡಿಯೋ ಮತ್ತು ಮನುಷ್ಯರ ಹಾವಭಾವಗಳನ್ನು (Gesticulations) ಏಕಕಾಲಕ್ಕೆ ಗ್ರಹಿಸಿ ಪ್ರತಿಕ್ರಿಯಿಸಲಿದೆ. ಉದಾಹರಣೆಗೆ, ಒಬ್ಬ ವ್ಯಕ್ತಿಯ ಮುಖದ ಭಾವನೆಯನ್ನು ನೋಡಿ ಆತ ಸಂತೋಷದಲ್ಲಿದ್ದಾನೋ ಅಥವಾ ಸಂಕಷ್ಟದಲ್ಲಿದ್ದಾನೋ ಎಂದು ಅರ್ಥೈಸಿಕೊಂಡು, ಅದಕ್ಕೆ ತಕ್ಕಂತೆ ಸಾಂತ್ವನದ ಭಾಷೆಯಲ್ಲಿ ಯಂತ್ರ ಸಂವಹನ ನಡೆಸಲಿದೆ.

ಧ್ವನಿ ಆಧಾರಿತ ಇಂಟರ್ನೆಟ್: ಇಂದಿನ ಇಂಟರ್ನೆಟ್ ಹೆಚ್ಚಾಗಿ ಅಕ್ಷರಸ್ಥರಿಗೆ ಮಾತ್ರ ಸೀಮಿತವಾಗಿದೆ. ಆದರೆ ಭವಿಷ್ಯದಲ್ಲಿ ಕೀಬೋರ್ಡ್‌ಗಳ ಅಗತ್ಯವೇ ಇರುವುದಿಲ್ಲ. ಗ್ರಾಮೀಣ ಭಾಗದ ಒಬ್ಬ ಸಾಮಾನ್ಯ ವ್ಯಕ್ತಿ ತನ್ನ ಆಡುಭಾಷೆಯಲ್ಲಿ ಇಂಟರ್ನೆಟ್ ಜೊತೆ ಮಾತನಾಡಿ ತನಗೆ ಬೇಕಾದ ಮಾಹಿತಿ ಅಥವಾ ಸರ್ಕಾರಿ ಸೌಲಭ್ಯಗಳನ್ನು ಪಡೆಯಬಹುದು. ಇದು 'ಡಿಜಿಟಲ್ ಸಾಕ್ಷರತೆ'ಯ ವ್ಯಾಖ್ಯಾನವನ್ನೇ ಬದಲಿಸಲಿದೆ.

ನೈಜ ಸಮಯದ ಸರ್ವಭಾಷಾ ಭಾಷಾಂತರ: ಭವಿಷ್ಯದಲ್ಲಿ ನಾವು ಧರಿಸುವ ಇಯರ್‌ಫೋನ್‌ಗಳು (Smart Earbuds) ಎದುರಿಗಿರುವ ವ್ಯಕ್ತಿ ಯಾವುದೇ ವಿದೇಶಿ ಭಾಷೆಯಲ್ಲಿ ಮಾತನಾಡಿದರೂ ಅದನ್ನು ತಕ್ಷಣವೇ ನಮ್ಮ ಮಾತೃಭಾಷೆಗೆ ತರ್ಜುಮೆ ಮಾಡಿ ನಮಗೆ ಕೇಳಿಸುತ್ತವೆ. ಇದು ಜಗತ್ತಿನ ಯಾವುದೇ ಮೂಲೆಯ ಇಬ್ಬರು ವ್ಯಕ್ತಿಗಳ ನಡುವಿನ ಭಾಷಾ ಗೋಡೆಯನ್ನು ಸಂಪೂರ್ಣವಾಗಿ ಕೆಡವಿ ಹಾಕಲಿದೆ.

ಉಪಸಂಹಾರ: ಅಸ್ಮಿತೆ ಮತ್ತು ತಂತ್ರಜ್ಞಾನದ ಸಮನ್ವಯ

ಭಾಷಾ ತಂತ್ರಜ್ಞಾನ ಮತ್ತು ಕೃತಕ ಬುದ್ಧಿಮತ್ತೆಗಳು ಮಾನವನ ಬುದ್ಧಿಶಕ್ತಿಗೆ

ಪರ್ಯಾಯವಲ್ಲ, ಬದಲಿಗೆ ಅವು ಮಾನವ ಸಾಮರ್ಥ್ಯವನ್ನು ವೃದ್ಧಿಸುವ 'ಪೂರಕ ಸಾಧನಗಳು'. ಈ ತಂತ್ರಜ್ಞಾನದ ಅತಿದೊಡ್ಡ ಕೊಡುಗೆ ಎಂದರೆ ಅದು ಭಾಷೆಗಳನ್ನು ಕೇವಲ ಸಂವಹನಕ್ಕಾಗಿ ಬಳಸುವುದಲ್ಲದೆ, ಅಳಿವಿನ ಅಂಚಿನಲ್ಲಿರುವ ಉಪಭಾಷೆಗಳನ್ನು ಮತ್ತು ಅಪಾರ ಸಾಹಿತ್ಯವನ್ನು ಡಿಜಿಟಲ್ ರೂಪದಲ್ಲಿ ಶಾಶ್ವತವಾಗಿ ಸಂರಕ್ಷಿಸಲು ನೆರವಾಗುತ್ತಿದೆ.

ಆದರೆ, ಈ ಹಾದಿಯಲ್ಲಿ ನಾವು ಎರಡು ಪ್ರಮುಖ ಅಂಶಗಳನ್ನು ನೆನಪಿಡಬೇಕು:

1. ಭಾಷಾ ಅಸ್ಮಿತೆಯ ರಕ್ಷಣೆ: ಯಂತ್ರಗಳು ನೀಡುವ ಭಾಷಾಂತರದ ಮೇಲೆ ಅತಿಯಾಗಿ ಅವಲಂಬಿತರಾಗಿ ನಮ್ಮ ಭಾಷೆಯ ಸೊಗಡು, ಶೈಲಿ ಮತ್ತು ಸ್ವಂತಿಕೆಯನ್ನು ನಾವು ಮರೆಯಬಾರದು. ತಂತ್ರಜ್ಞಾನವು ಭಾಷೆಯನ್ನು ಬಳಸಲು ಸಹಾಯ ಮಾಡಬೇಕೇ ಹೊರತು ಭಾಷೆಯ ಮೌಲ್ಯವನ್ನು ಕುಗ್ಗಿಸಬಾರದು.
2. ನೈತಿಕ ಬಳಕೆ (Ethical AI): ಸುಳ್ಳು ಸುದ್ದಿಗಳನ್ನು ಸೃಷ್ಟಿಸಲು ಅಥವಾ ಸಮಾಜದಲ್ಲಿ ದ್ವೇಷ ಹರಡಲು ಈ ತಂತ್ರಜ್ಞಾನ ಬಳಕೆಯಾಗದಂತೆ ನೋಡಿಕೊಳ್ಳುವುದು ನಮ್ಮ ಜವಾಬ್ದಾರಿಯಾಗಿದೆ.

ತಂತ್ರಜ್ಞಾನವು ಮನುಷ್ಯನಿಗೆ ಸಿಕ್ಕಿರುವ ವರ. ಇದನ್ನು ವಿವೇಚನೆಯಿಂದ ಬಳಸಿದರೆ, ಕನ್ನಡದಂತಹ ಪ್ರಾದೇಶಿಕ ಭಾಷೆಗಳು ಕೇವಲ ಮನೆಯ ಅಂಗಳಕ್ಕೆ ಸೀಮಿತವಾಗದೆ ಜಾಗತಿಕ ವೇದಿಕೆಯಲ್ಲಿ ಹೆಮ್ಮೆಯಿಂದ ಮಿನುಗಲಿವೆ. ಮನುಷ್ಯನ ಸೃಜನಶೀಲತೆ ಮತ್ತು ಯಂತ್ರದ ವೇಗ ಒಟ್ಟಾಗಿ ಸಾಗಿದಾಗ ಮಾತ್ರ ಒಂದು ಸುಂದರ ಮತ್ತು ಭಾಷಾ ಅಡೆತಡೆಗಳಿಲ್ಲದ ಭವಿಷ್ಯ ನಿರ್ಮಾಣವಾಗಲು ಸಾಧ್ಯ.

ಪರಾಮರ್ಶನ ಗ್ರಂಥಗಳು:

1. ಜರಫ್ಸ್ಕಿ , ಡಿ. ಮತ್ತು ಮಾರ್ಟಿನ್, ಜೆ. ಹೆಚ್. (2023): Speech and Language Processing. (NLP ಕ್ಷೇತ್ರದ ಅತ್ಯಂತ ಪ್ರಮಾಣಿತ ಗ್ರಂಥ).
2. ರಸೆಲ್, ಎಸ್. ಮತ್ತು ನಾರ್ವಿಗ್, ಪಿ. (2020): Artificial Intelligence: A Modern Approach. (AI ನ ಮೂಲಭೂತ ಸಿದ್ಧಾಂತಗಳನ್ನು ತಿಳಿಯಲು ಉಪಯುಕ್ತ).
3. ಬರ್ಡ್, ಎಸ್. ಮತ್ತು ಇತರರು (2009): Natural Language Processing with Python. (ಪೈಥಾನ್ ಬಳಸಿ ಭಾಷಾ ತಂತ್ರಜ್ಞಾನ ರೂಪಿಸುವ ಮಾಹಿತಿ).
4. ಸಂಚಯ (Sanchaya) ಆನ್‌ಲೈನ್ ಪೋರ್ಟಲ್: ಕನ್ನಡ ಭಾಷಾ ದತ್ತಾಂಶ ಮತ್ತು ತಂತ್ರಜ್ಞಾನದ ಮೇಲಿನ ಸಂಶೋಧನಾ ಪ್ರಬಂಧಗಳು.
5. ಮನ್ನಿಂಗ್, ಸಿ. ಡಿ. ಮತ್ತು ಶೂಟ್ಜೆ, ಹೆಚ್. (1999): Foundations of Statistical Natural Language Processing. (ಸಂಖ್ಯಾಶಾಸ್ತ್ರೀಯ ಆಧಾರಿತ NLP ಮಾದರಿಗಳನ್ನು

ಅರ್ಥಮಾಡಿಕೊಳ್ಳಲು ಅತ್ಯಂತ ಪ್ರಮುಖವಾದ ಪುಸ್ತಕ).

6. ಗುಡ್‌ಫೆಲೋ, ಐ. ಮತ್ತು ಇತರರು (2016): Deep Learning. (ನ್ಯೂರಲ್ ನೆಟ್‌ವರ್ಕ್‌ಗಳು ಮತ್ತು ಡೀಪ್ ಲರ್ನಿಂಗ್ ತಂತ್ರಜ್ಞಾನಗಳು ಭಾಷೆಯನ್ನು ಹೇಗೆ ವಿಶ್ಲೇಷಿಸುತ್ತವೆ ಎಂಬುದರ ಬಗ್ಗೆ ಇದು ಸಮಗ್ರ ಮಾಹಿತಿ ನೀಡುತ್ತದೆ).

7. ಬಿ. ಕೆ. ಶರ್ಮ (2021): Natural Language Processing. (ಭಾರತೀಯ ಸಂದರ್ಭದಲ್ಲಿ ಮತ್ತು ಭಾರತೀಯ ಭಾಷೆಗಳಿಗೆ ಅನ್ವಯವಾಗುವಂತೆ NLP ಸಿದ್ಧಾಂತಗಳನ್ನು ವಿವರಿಸುವ ಪುಸ್ತಕ).

8. ಕೆ. ಪಿ. ರಾವ್ ಅವರ ಲೇಖನಗಳು: ಕನ್ನಡದ ಕೀಬೋರ್ಡ್ ವಿನ್ಯಾಸ ಮತ್ತು ಕನ್ನಡ ಕಂಪ್ಯೂಟಿಂಗ್ ಬಗ್ಗೆ ಇವರು ಮಾಡಿರುವ ಮೂಲಭೂತ ಸಂಶೋಧನೆಗಳು ಅತ್ಯಂತ ಮೌಲ್ಯಯುತವಾದವು.

9. ಹಂಪಿ ಕನ್ನಡ ವಿಶ್ವವಿದ್ಯಾಲಯದ ಪ್ರಕಟಣೆಗಳು: 'ಕಂಪ್ಯೂಟರ್ ಮತ್ತು ಕನ್ನಡ' ಎಂಬ ಶೀರ್ಷಿಕೆಯಡಿ ಪ್ರಕಟವಾಗಿರುವ ಸಂಶೋಧನಾ ಸಂಪುಟಗಳು ಕನ್ನಡ ಭಾಷಾ ತಂತ್ರಜ್ಞಾನದ ಬೆಳವಣಿಗೆಯ ಇತಿಹಾಸವನ್ನು ದಾಖಲಿಸಿವೆ.

10. ಸಿ–ಡಾಕ್ (C–DAC) ಪ್ರಕಟಣೆಗಳು: ಭಾರತೀಯ ಭಾಷೆಗಳಿಗಾಗಿ ಜಿ–ಪಿಸ್ಟ್ (GIST) ತಂತ್ರಜ್ಞಾನ ಮತ್ತು ಅಕ್ಷರ ವಿಶ್ಲೇಷಣೆ ಕುರಿತಾದ ಸಂಶೋಧನಾ ವರದಿಗಳು.

11. TDIL (Technology Development for Indian Languages) ಪೋರ್ಟಲ್: ಭಾರತ ಸರ್ಕಾರದ ಎಲೆಕ್ಟ್ರಾನಿಕ್ಸ್ ಮತ್ತು ಮಾಹಿತಿ ತಂತ್ರಜ್ಞಾನ ಸಚಿವಾಲಯದ ಅಧಿಕೃತ ಪೋರ್ಟಲ್. ಇಲ್ಲಿ ಕನ್ನಡ ಸೇರಿದಂತೆ ವಿವಿಧ ಭಾರತೀಯ ಭಾಷೆಗಳ ತಾಂತ್ರಿಕ ಅಭಿವೃದ್ಧಿಯ ವರದಿಗಳು ಲಭ್ಯವಿವೆ.

12. ಭಾಷಿಣಿ (Bhashini – Digital India): ರಾಷ್ಟ್ರೀಯ ಭಾಷಾ ಅನುವಾದ ಮಿಷನ್‌ನ ಅಧಿಕೃತ ವೆಬ್‌ಸೈಟ್. ಇದರಲ್ಲಿ AI ಬಳಸಿಕೊಂಡು ಭಾರತೀಯ ಭಾಷೆಗಳ ನಡುವಿನ ಅಂತರವನ್ನು ಕಡಿಮೆ ಮಾಡುವ ಪ್ರಯೋಗಗಳ ಬಗ್ಗೆ ಹೊಸ ದತ್ತಾಂಶಗಳಿವೆ.

Funding:

This study was not funded by any grant.

Conflict of interest:

The Authors have no conflict of interest to declare that they are relevant to the content of this article.

About the License:

© The Authors 2024. The text of this article is open access and licensed under a Creative Commons Attribution 4.0 International License.