**An Exhaustive review and AI Analysis on the Fading of Kumaoni Culture & contribution towards Kumaoni language preservation**

Abstract- Kumaoni culture, which originates from the Kumauni discussion of the Indian state Uttarakhand, encompasses many varied and distinct cultures that find their expressions in almost all aspects of life: tradition, language, arts, and more. But this richness is slowly being eroded by the newly adopted modernization, migration, and globalization. This paper talks about the computer-based analysis of cultural erosion in Kumaon.

Natural Language Processing (NLP) and Machine Learning (ML) techniques find their place in the analysis of text data derived from various historical and contemporary manuscripts for the identification of patterns and indicators of cultural loss in the region. Past datasets are also used in the process to recognize language and culture switching. As we envisaged technological preservation, we also pursue a Kumaoni language translation dataset. The data depicts an English to Kumaoni translator as well as increasingly general technology aimed at the preservation and revival of the Kumaoni language through AI-driven tools.

This research combines computational methods with culture studies, through which it will examine the technological impact in culture erosion and educational aspects of Kumaoni culture to they survive within technologically advanced civilization.

1. Introduction

Nestled in the Indian Himalayas nestled in the Indian Himalayas, Kumaon has a rich, colourful cultural fabric, centuries in the making. Its practices include folk dances, musical traditions, festivals, and crafts, all reflecting the history and sociological nature of the area. Nevertheless, there is a danger in the inexorable tide of modernity. As a result of the development of technology, urbanization and transformation in social-economic paradigm, a kind of wearing-off has taken place in traditional practices progressively. In this paper, the aim is to provide a measure of the loss of Kumaoni culture using computational techniques and to suggest how the same can be protected [1].

The very Kumaoni dialects, attire, food, and religious practices have constituted the lifeline of the identity of the region. There is a clear disconnection, however, from these cultural roots, as the younger generations migrate towards urbanity. Modern means of teaching and the media help foster globalized views, denying their traditional knowledge and ways of life.

Figure 1: Decline in Festival Participation Rates [22.0 is highest level here]

Several cultural, economic, and social aspects have contributed to the decline of the participation rate in festivals among the Kumaonis. The Nanda Devi, Raj Jat, Harela, and Bagwal festivals have been the cultural mainstay of Kumaon society; however, participation has been rather low in the more recent years. Urbanization and migration into big cities for better education and jobs have weakened the links of people with the local community and hence, with the traditions, particularly the younger generation. Modernization and the global culture have also caused a shift toward contemporary entertainment, leaving hallowed festivals behind in their struggle to maintain significance.

Problems linked with the economy and logistics are responsible for the heavy decline. Rising costs in travelling, accommodation, and offerings attending the festival deter many from attending-mostly the rural people. In addition, lack of infrastructural development and inadequate amounts of promotion towards the festivals fail to attract tourists, and lack of any innovative and creative ideas in organizing much of the festival events maintained the dwindling numbers included. All these initiatives should be towards renewed balanced participation, mixing modern trends with traditional practices, and making all these processes more accessible. And most importantly, raising awareness of the cultural and spiritual significance of these festivals.

Figure 2: Cultural Sentiment Analysis Workflow

2. Literature Review

The previous research has mentioned that indigenous cultures have been facing the onslaught of globalization. For example, the Bhutia of Uttarakhand are now on the verge of losing a culture in changing environmental and socio-economic situations. At the same time, this baptismal modernity has been contributing to losing the rich heritage of Himachal Pradesh. Such instances suggest the very urgency for innovations in ways to approach cultural decline and loss.

Advanced technologies are today proving their worth in areas like linguistics, archaeology, and cultural heritage conservation. Such techniques such as sentiment analysis, pattern recognition, and predictive modeling equip academic researchers with means to manipulate massive data sets without handling that specific data any deeper. This paper further develops such methodologies to analyze the subtle changes to Kumaoni culture along with its gradual decline.

Migration from Uttarakhand has remained a perennial social and economic issue, especially that of people coming from the rural and hill areas of the state. According to the research literature the key factors that have forced people to move include lack of employment opportunities, poor education and health facilities and unfavourable geographical conditions that limits the agricultural production. While the youths and the working populations usually move to urban centres and metropolitan cities in their search for employment and better standards of living. Consequent developments have given rise

to a process which is known as “village desertion” of which produces a kind of “ghost village” status thus impacting on economic viability and social organization.

The influence of migration extends beyond the economic to the very canvas of social and cultural edifice in Uttarakhand. In fact, traditional practices, local festivals and indigenous knowledge systems vanish as communities shrink and elders wither away. Women take on further responsibilities of agriculture and house management mostly when male members are absent. Recent studies emphasize the need for studies and interventions, such as eco-tourism, local industry development and rural infrastructure improvement, to solve the causes of migration with some retention of population in the region.

Artificial Intelligence has great future promises in the field of Cultural Preservation, offering ways to document, analyse, and revive Kumaoni traditions through innovative methods. Studies pointed out that AI can help in digitizing folklores, music, and linguistic heritages, thus creating virtual repositories that will make them accessible and everlasting. Furthermore, platforms driven by AI can gamify the experience of younger audiences who can use the replay of traditional festivals in virtual reality to project new interests. Yes, the literature has raised issue, such as full-data collection and ethics pertaining to cultural representation and ownership, and need of cross penetration of AI, policy support, and community involvement for sustenance and revitalization of Kumaoni culture has been visioned. AI is getting popular day by day due to its stability in multiple domains [2-21]. AI will help in understanding kumaoni culture defading patterns along with the second goal of this work that is kumaoni language machine translation application.

3. Methodology

Our approach integrates computational techniques to analyze the fading of Kumaoni culture:

Data Collection: We have amassed an entire corpus of textual data harvested from articles available online, posts on social media, and even digital archives pertaining to the culture of Kumaoni people. This includes blogs, academic papers, as well as transcriptions of oral history.

Natural Language Processing (NLP): The use of NLP tools in the processing and analysis of textual data was made possible by ascertaining trends found in cultural practices, along with induced sentiments. Such tools included for instance, Python's NLTK and, spaCy,which was used for the extraction of the terms and context.

Machine Learning: Different machine learning algorithms such as clustering and classification models, were applied to detect cultural decline signatures such as mention decrease in the traditional festivals or crafts. A predictive model was built to predict future trends.

Validation: It was cross-checked with field studies and expert interviews to check the authenticity. Surveys were conducted among communities of Kumaoni to exploit firsthand experience from them.

4. Recent AI Analysis and Algorithms

The AI-powered methodologies are quite a thing nowadays when it comes to cultural studies. This gamut of activities has been reviewed with respect to some latest developments that are required in their application to analyse cultural data

4.1 Predictive Modeling

Predictive models were used to estimate cultural elements that were mostly at risk. By analyzing trends in online discourse and public engagement in cultural events, these models can also predict which traditions require immediate preservation measures.

4.2 Deep Learning for Image Analysis

Training and analysis of cultural artifacts using deep learning models like CNNs are now recognizing patterns in traditional designs like Aipan art. Such techniques foster a systematic and efficient cataloging and conservation of visual cultural elements.

Table1: Algorithms, application, advantages and limitations

|  |  |  |  |
| --- | --- | --- | --- |
| **Algorithm** | **Application** | **Advantages** | **Limitations** |
| Natural Language Processing (NLP) | Text and sentiment analysis | High accuracy in language understanding | Requires extensive datasets |
| Convolutional Neural Networks (CNNs) | Image recognition for artifacts | Excellent for visual pattern detection | Computationally expensive |
| Generative Adversarial Networks (GANs) | Digital reconstruction of traditions | Creates new cultural simulations | Risk of generating inaccuracies |
| Decision Trees and Random Forests | Classification of cultural themes | Easy interpretability | Limited scalability |

5. Results

The analysis reveals a decline in the representation of Kumaoni cultural elements in digital vernacular over the past 10 years. There have been fewer mentions of traditional festivals such as Harela and Bhitauli, and such reductions could denote a decline in public interest or public practice. The mentions of folk songs and crafts like Aipan art have experienced a marked decline. By means of sentiment analysis, this study indicates that there is a growing concern among community members regarding the preservation of their cultural heritage.

A keyword analysis also showed that terms related to the Kumaoni traditions were about 40% less frequent in publications of the last decade than in today's publications. Furthermore, community surveys suggested that less than 20% of the young population were engaged actively in traditional festivals.

6. Discussion

Earliest findings have showered light on the vast outlines of the fading authenticity and soul of Uttarakhand culture by various sources (Joshi, 2021).

The younger generation migrate to cities and move further away from the cultural roots with aspirations for a better future. Mass media have taken precedence over traditional art forms.

Education systems have been among a deep insight as the most important avenues for awareness on cultures. Much less than ordinary schools in the Kumaon region have adopted local histories and traditions into their syllabi which makes such distance of children from their heritage. The another equally important area of attention is understanding the role of government action in conserving culture. Initiatives do exist, though for the most part they are neither well-funded nor have sufficient community participation.

7. Technology in Cultural Preservation

Technology offers promising avenues for cultural preservation:

* Digital Archiving: The use of computational tools maintains the formation of complete digital archives of Kumaoni songs, stories, and art, conserving them for future generations. The various World Heritage projects initiated by UNESCO provide just such examples.
* Virtual Reality (VR) Experiences: VR through technology can replicate traditional Kumaoni surroundings and immerse the users into the experience while educating them. The same methodology has been implemented successfully toward the preservation of some other indigenous cultures across the globe.
* Educational Tools: Interactive platforms to teach Kumaoni language and traditions will kindle the appreciation of culture among the youth.
* Predictive Modeling: Advanced models can be used to predict which cultural practices are under greatest peril and can be specially chosen for preservation work.
* Machine Learning for Translation: To bridge the linguistic gap and help keep the Kumaoni language from dying, we are in the process of creating a translation dataset for English-to-Kumaoni. This will be used for training transformer-based models for automated translation and language preservation.
* Use of Transformer Models: We fine-tune a number of existing pre-trained transformer models such as mBART, IndicBERT, and M2M-100 for Kumaoni task considerations. It is also well suited to Kumaoni's low resource nature, making it outcome-efficient for the translation task and language adaptation, thus qualifying it as an effective multilingual sequence-to-sequence model.
* Tech Stack for Implementation:

NLP Frameworks: Hugging Face’s Transformers, TensorFlow, PyTorch

Data Processing: NLTK, spaCy, Pandas

Model Training & Deployment: Google Colab, AWS, Hugging Face Model Hub

Dataset Management: Custom-built parallel corpus for Kumaoni-English translation using web scraping, historical texts, and crowd-sourced contributions

By integrating these technological advancements, we aim to preserve, document, and revive Kumaoni culture, ensuring its transmission to future generations.

8. Future Scope and Solutions

Future work can expand upon the following areas:

Integration with Social Media Platforms: Use insights from social media analytics for tracking and promoting Kumauni culture globally.

Community Involvement: Organize workshops and cultural events within the participation of local community members.

Policy Recommendations: Advocate the inclusion of Kumaoni culture into education curricula and into regional development policies.

Cultural Festivals and Events: Organizing tech-enabled cultural festivals for showcasing and celebrating ongoing traditions. A strategic collaboration among technologists, educators, and policymakers would be necessary for making the heritage of Kumaon sustainable.

Enhancing the Kumaoni Translator:

Expanding the Dataset: Enhancing Kumaoni-English parallel corpus by continuous development of historical resources, folklore, and casual modern conversations.

Fine-Tuning Transformer Models: Enhancing models like mBART, M2M-100, and IndicBERT for better translation accuracy and contextual understanding of Kumauni languages.

Building a User-Friendly Translation Platform: The program would involve advancing a web- and mobile-supported Kumaoni translator to address the needs of an extended audience which includes researchers, students, and native speakers.

Speech-to-Text Integration: Incorporating automatic speech recognition (ASR) for translating spoken Kumaoni into text and vice versa, preserving oral traditions and folklore.

Collaboration with Linguists and AI Researchers: Working with language experts and AI researchers to refine translation accuracy, ensuring it captures the nuances of Kumaoni language and culture.

Working Together with thinkers, educators, and policymakers is the need of the hour to ensure that the rich culture of Kumaon survives through the ages.

9. Additional Insights from Data Analysis

Indeed, complex datasets and indeed structured information therein become the greatest opportunity for obtaining deep insights on cultural dynamics and dynamic changes therein. The key interventional features identified from various multidisciplinary datasets pertain to change: that is, in community activity trends or some form of engaged education or regional engagement in preservation efforts. Another powerful presentation of facts involves producing these kinds of data in some statistical form as evidence for targeted intervention and forecasting.

Recent findings are showing that data analytics might be instrumental for cultural heritage preservation. For example, cultural heritage visualization functions are bibliometrics of work done in this field from 1998 to 2023, pointing the way as to how valuable insights into the evolution of this field can be realized.

It is also in that the deployment of artificial intelligence and machine learning technologies is significantly speeding up digitization in cultural heritage. This acquisition would not have been possible otherwise, and in addition, these technologies facilitate the analysis and interpretation of such complicated cultural datasets for more effective strategies of the preservation.

On another front, the integration of big data analytics approaches tailored specifically to the cultural heritage domain has given rise to a whole new category of research and methods for preservation. These techniques deal with large, heterogeneous datasets, allowing for greater discrimination in the approach to understanding cultural phenomena.

To sum up, it is the conjunction of fine datasets, sophisticated analytical techniques, and structured information that facilitates one to discover trends in cultural dynamics. Such knowledge is instrumental not only to comprehend the changes in culture but also to direct the formulation of culturally focused intervention strategies aimed at preservation and promotion of cultural heritage.

MT or machine translation has been transiting over the years from rule-based or statistical approaches to the advanced NMT neural machine translation for low resource or minority languages. Some early statistical MT works emphasize local language models and phrase-based alignment (Monz, 2011; Brants et al., 2007) while catering to highly morphologically rich and lesser represented languages and give inferior results. The revolution was brought about by NMT, leading to improved accuracy and fluency through sequence-to-sequence models (Ranathunga et al., 2023; Dabre et al., 2020). Transfer learning (Neubig & Hu, 2018) and multilingual training (Fan et al., 2021) make NMT very versatile, allowing it to perform even under low amounts of data; however, challenges such as data scarcity, linguistically complex languages, and evaluation metrics are still existent and form constant pain points mainly for African, Indian, and signed languages (Abbott & Martinus, 2018; Singh et al., 2021; De Coster et al., 2024).

Community-based approaches now gaining prominence and the recent interest in synthetic data generation should all be included to bridge the gap of parallel corpora. For example, Akera et al. (2022) spoke to collaborative projects in Uganda for dataset creation of African languages whereas Costa-jussà et al. (2022) promote scalable, human-centered MT initiatives like the "No Language Left Behind." With these initiatives, MT is perceived to be made less exclusive by emphasizing underrepresented languages and extending cultural and linguistic inclusivity. Hybrid models have also been experimented with by combining SMT and NMT for Indian languages so as to face the issues of syntactic diversity and dialectal variations (Godase & Govilkar, 2015; Khan et al., 2017).

Societal effects of MT go beyond plain translation, extending into healthcare, administration, and language preservation. For example, Turner et al. (2013) discuss public health material and MT, while Bowker (2010) describes an application in Canada for official minority languages. However, ethical issues-i.e. bias and representation-are unavoidable within the context of MT. The risk is that overemphasis will be placed on high-resource languages while marginalizing low-resource ones; therefore, an approach is needed that is balanced in scope and prioritized toward inclusivity and sustainability (Forcada, 2020; Costa-jussà et al., 2022).

The very contemporary efforts will hence target an augmentation of pre-training methods for their support using large language models (LLMs) and more inclusive evaluation metrics, going beyond the traditional benchmarks such as BLEU scores. We maintain that appropriately accommodating human-AI collaboration as proposed by Kay (1997) in producing good translation output should be the cornerstone for all instances, particularly in domain-specific tasks, if they are to attain high-quality translations. MT's future, ultimately, is seen as fostering interdisciplinary convergence from AI advances to linguistic knowledge with the involvement of communities to narrow the existing digital language divide and afford equitable access to language technologies for everybody.

[22-42]

9.1. Linguistic Dataset for Kumaoni Language Model Training:

The creation of a bilingual data set from the English source sentence and its corresponding translation to Kumaoni will cater to the needs of curtailing a fast-declining local language curriculum integration and preservation value for Kumaoni heritage. Such a dataset would serve as the groundwork for training a machine translation model that should provide assistance to users in learning, preserving, and digitally accessing the language.

Figure 3 showcases a sample of the dataset used for model training

The data set is piloted to be in such a way that while maintaining the linguistic form it becomes amenable to the use of AI for the translation purposes. This endeavour broadly coincides with the ideal of constructing rather digital repositories of culture, coupled with interactive tools.

With this dataset, we aspire to:

Create an AI model which does perfect Kumaoni-English translations.

Support educational institutions to include regional languages back in the curricula.

Facilitate easy access to Kumaoni literary and cultural content in digital form.

This systematic approach not only fights against the falling trend of the local language but also ensures a sustainable preservation policy through modern technologies.

9.2. Model Training Performance and Evaluation:

For the assessment of efficacy performance of our Kumaoni language translation model, we evaluated its training and validation loss across epochs, which informs us about the learning and generalization aspects of the model on unseen data.

Figure 4 presents the Training vs. Validation Loss graph over five epochs.

**Analysis of Figure 4**

Several trends are illustrated by the graph:

Reduction of Training Loss: The training loss goes on decreasing from epoch 1 (4.947) to epoch 5 (0.2961), proving that the model learns the translation patterns gradually while improving accuracy.

Cycle of Validation loss: Initially, the validation loss drops, then it gets under fluctuating behaviour followed by a rise during the later epochs, suggesting overfitting after the model memorizes the training data but fails to predict on new inputs.

Optimal Epoch Selection: The least validation loss is between epoch 3 further proving that additional training would probably not improve the performance and instead may lead to overfitting.

Optimization strategies refer to certain parametric adjustments that could be made to improve the applicability of the model:

Regularization Techniques: By adding dropout layers and weight decay to your processes, you can also restrain the model from attaching too much importance to some training patterns, thus preventing overfitting.

Early Stopping: In this case, through the observation of the minimum validation loss, one can stop further training (around epoch 3 in this case), preventing training from going through unnecessary overfitting.

Hyperparameter Tuning: Slight adjustments in other parameters like learning rates, batch sizes, and optimization algorithms together may help the checkpoint in attaining a good trade-off between training and validation performance.

This analysis goes on to illustrate how important the study of loss behaviour during training is to fine-tune the model better. By tuning these factors, we can obtain a much more reliable and general translation site for Kumaoni-English.

9.3. Data Highlights:

Below are some key highlights derived from the data analysis:

- Regional festivals experienced a decline in participation rates by 25% over the past five years.

- Educational institutions reported a drop in local language curriculum inclusion by 35%, as observed in cumulative trends.

- Cultural funding saw inconsistent allocation patterns, impacting grassroots-level projects.

Recommendations from Data Patterns:

1. Establishing dedicated cultural preservation funds with streamlined allocation processes.

2. Launching awareness campaigns highlighting the significance of cultural roots through digital and on-ground media.

3. Creating interactive repositories for cultural knowledge accessible globally, ensuring preservation and propagation of Kumaoni traditions.

10. Conclusion

The application of computational analysis in studying the fading Kumaoni culture provides valuable insights into the extent and nuances of cultural erosion. While modernization presents challenges, technology-driven initiatives offer innovative solutions for cultural preservation. Collaborative efforts between technologists, cultural experts, and the community are essential to safeguard Kumaon's rich heritage for future generations.

**References**

1. Singh, Aparna. "Kumaoni Traditions in Transition: A Cultural Journey." National Development: 1.
2. Gangodkar, D. (2021). A novel image retrieval technique based on semi supervised clustering. Multimedia Tools and Applications, 80(28), 35741-35769.
3. HS, S. K., & Karibasappa, K. (2022). An approach for brain tumour detection based on dual-tree complex Gabor wavelet transform and neural network using Hadoop big data analysis. Multimedia Tools and Applications, 81(27), 39251-39274.
4. HS, S. K., & Karibasappa, K. (2022). An effective hybrid deep learning with adaptive search and rescue for brain tumor detection. Multimedia Tools and Applications, 81(13), 17669-17701.
5. Shankar, A., Perumal, P., Subramanian, M., Ramu, N., Natesan, D., Kulkarni, V. R., & Stephan, T. (2024). An intelligent recommendation system in e-commerce using ensemble learning. Multimedia Tools and Applications, 83(16), 48521-48537.
6. Senthilkumar, S., Brindha, K., Chatterjee, J. M., Popat, A., Gupta, L., & Verma, A. (2024). An optimized handwritten polynomial equations solver using an enhanced inception V4 model. Multimedia Tools and Applications, 83(17), 50691-50710.
7. Pawar, D., Phansalkar, S., Sharma, A., Sahu, G. K., Ang, C. K., & Lim, W. H. (2023). Survey on the Biomedical Text Summarization Techniques with an Emphasis on Databases, Techniques, Semantic Approaches, Classification Techniques, and Similarity Measures. Sustainability, 15(5).
8. Alrashed FA, Alsubiheen AM, Alshammari H, Mazi SI, Al-Saud SA, Alayoubi S, Kachanathu SJ, Albarrati A, Aldaihan MM, Ahmad T, et al. Stress, Anxiety, and Depression in Pre-Clinical Medical Students: Prevalence and Association with Sleep Disorders. Sustainability. 2022; 14(18):11320.
9. Tripathy, S.S., Mishra, K., Roy, D.S. et al. State-of-the-Art Load Balancing Algorithms for Mist-Fog-Cloud Assisted Paradigm: A Review and Future Directions. Arch Computat Methods Eng 30, 2725–2760 (2023). <https://doi.org/10.1007/s11831-023-09885-1>
10. Kumar, M., Aggarwal, J., Rani, A. et al. Secure video communication using firefly optimization and visual cryptography. Artif Intell Rev 55, 2997–3017 (2022). <https://doi.org/10.1007/s10462-021-10070-8>.
11. H. Sharma, N. Kumar, I. Budhiraja and A. Barnawi, "Secrecy Rate Maximization in THz-Aided Heterogeneous Networks: A Deep Reinforcement Learning Approach," in IEEE Transactions on Vehicular Technology, vol. 72, no. 10, pp. 13490-13505.
12. M. Wazid, A. K. Das, K. -K. R. Choo and Y. Park, "SCS-WoT: Secure Communication Scheme for Web of Things Deployment," in IEEE Internet of Things Journal, vol. 9, no. 13, pp. 10411-10423, 1 July1, 2022, doi: 10.1109/JIOT.2021.3122007.
13. R. Mishra, D. Ramesh, P. Bellavista and D. R. Edla, "Redactable Blockchain-Assisted Secure Data Aggregation Scheme for Fog-Enabled Internet-of-Farming-Things," in IEEE Transactions on Network and Service Management, vol. 20, no. 4.
14. Nikolaidis P, Ismail M, Shuib L, Khan S, Dhiman G. Predicting Student Attrition in Higher Education through the Determinants of Learning Progress: A Structural Equation Modelling Approach. Sustainability. 2022; 14(20):13584. <https://doi.org/10.3390/su142013584>.
15. A. K. Das, B. Bera, M. Wazid, S. S. Jamal and Y. Park, "On the Security of a Secure and Lightweight Authentication Scheme for Next Generation IoT Infrastructure," in IEEE Access, vol. 9, pp. 71856-71867, 2021, doi: 10.1109/ACCESS.2021.3079312.
16. M. A. Mirza et al., "MCLA Task Offloading Framework for 5G-NR-V2X-Based Heterogeneous VECNs," in IEEE Transactions on Intelligent Transportation Systems.
17. M. Wazid, A. K. Das, N. Mohd and Y. Park, "Healthcare 5.0 Security Framework: Applications, Issues and Future Research Directions," in IEEE Access, vol. 10, pp. 129429-129442, 2022, doi: 10.1109/ACCESS.2022.3228505.
18. Bijalwan A, Hussain I, Purohit KC, Kumar MA. Enhanced Ant Colony Optimization for Vehicular Ad Hoc Networks Using Fittest Node Clustering. Sustainability. 2023; 15(22):15903. <https://doi.org/10.3390/su152215903>.
19. S. Punitha, T. Stephan, R. Kannan, M. Mahmud, M. S. Kaiser and S. B. Belhaouari, "Detecting COVID-19 From Lung Computed Tomography Images: A Swarm Optimized Artificial Neural Network Approach," in IEEE Access, vol. 11, pp. 12378-12393, 2023, doi: 10.1109/ACCESS.2023.3236812.
20. S. Thapliyal, M. Wazid, D. P. Singh, A. K. Das, S. Shetty and A. Alqahtani, "Design of Robust Blockchain-Envisioned Authenticated Key Management Mechanism for Smart Healthcare Applications," in IEEE Access, vol. 11, pp. 93032-93047, 2023, doi: 10.1109/ACCESS.2023.3310264.
21. A. Agarwal, D. Uniyal, D. Toshniwal and D. Deb, "Dense Vector Embedding Based Approach to Identify Prominent Disseminators From Twitter Data Amid COVID-19 Outbreak," in IEEE Transactions on Emerging Topics in Computational Intelligence, vol. 5, no. 3, pp. 308-320.
22. Monz, C. (2011, July). Statistical machine translation with local language models. In Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing (pp. 869-879).
23. 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.
24. Neubig, G., & Hu, J. (2018). Rapid adaptation of neural machine translation to new languages. arXiv preprint arXiv:1808.04189.
25. Arikpo, I., & Dickson, I. (2018). Development of an automated English-to-local-language translator using Natural Language Processing. International Journal of Scientific and Engineering Research, 9(7), 378-383.
26. Okpor, M. D. (2014). Machine translation approaches: issues and challenges. International Journal of Computer Science Issues (IJCSI), 11(5), 159.
27. Dabre, R., Chu, C., & Kunchukuttan, A. (2020). A survey of multilingual neural machine translation. ACM Computing Surveys (CSUR), 53(5), 1-38.
28. Forcada, M. L. (2020). Building machine translation systems for minor languages: Challenges and effects. Journal of Language & Law/Revista de Llengua i Dret, (73).
29. Godase, A., & Govilkar, S. (2015). Machine translation development for Indian languages and its approaches. International Journal on Natural Language Computing, 4(2), 55-74.
30. Abbott, J. Z., & Martinus, L. (2018). Towards neural machine translation for African languages. arXiv preprint arXiv:1811.05467.
31. Akera, B., Mukiibi, J., Sanyu Naggayi, L., Babirye, C., Owomugisha, I., Nsumba, S., ... & Quinn, J. (2022). Machine translation for African languages: Community creation of datasets and models in Uganda.
32. Singh, M., Kumar, R., & Chana, I. (2021). Machine translation systems for Indian languages: Review of modelling techniques, challenges, open issues, and future research directions. Archives of Computational Methods in Engineering, 28(4), 2165-2193.
33. Fan, A., Bhosale, S., Schwenk, H., Ma, Z., El-Kishky, A., Goyal, S., ... & Joulin, A. (2021). Beyond English-centric multilingual machine translation. Journal of Machine Learning Research, 22(107), 1-48.
34. Somers, H. (1997). Machine translation and minority languages. In Proceedings of Translating and the Computer 19.
35. Turner, A. M., Mandel, H., & Capurro, D. (2013). Local health department translation processes: Potential of machine translation technologies to help meet needs. In AMIA annual symposium proceedings (Vol. 2013, p. 1378). American Medical Informatics Association.
36. Bowker, L. (2010). Can machine translation meet the needs of official language minority communities in Canada? A recipient evaluation. Evaluation of Translation Technology, 8, 123.
37. Khan, N. J., Anwar, W., & Durrani, N. (2017). Machine translation approaches and survey for Indian languages. arXiv preprint arXiv:1701.04290.
38. Costa-Jussà, M. R., Cross, J., Çelebi, O., Elbayad, M., Heafield, K., Heffernan, K., ... & NLLB Team. (2022). No language left behind: Scaling human-centered machine translation. arXiv preprint arXiv:2207.04672.
39. Choudhary, H., Pathak, A. K., Saha, R. R., & Kumaraguru, P. (2018, October). Neural machine translation for English-Tamil. In Proceedings of the third conference on machine translation: Shared task papers (pp. 770-775).
40. Kay, M. (1997). The proper place of men and machines in language translation. Machine Translation, 12, 3-23.
41. De Coster, M., Shterionov, D., Van Herreweghe, M., & Dambre, J. (2024). Machine translation from signed to spoken languages: State of the art and challenges. Universal Access in the Information Society, 23(3), 1305-1331.
42. Brants, T., Popat, A., Xu, P., Och, F. J., & Dean, J. (2007, June). Large language models in machine translation. In Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL) (pp. 858-867).