International Journal of Engineering Trends and Applications (IJETA) – Volume 13 Issue 3, June 2026
RESEARCH ARTICLE OPEN

Prompt-Based Multilingual Summarization in Hindi and Marathi Using Large Language Models

Prof. Wrushabh Shirsat, Sakshi Sarode, Nilesh Khedkar, Saad Kazi, Om Sakunde
Department of Information Technology,
Zeal College of Engineering and Research, SPPU, Pune
Pune (MS), India

ABSTRACT
The exponential growth of multilingual digital content has created an increasing demand for intelligent systems capable of extracting and presenting concise information efficiently. Automatic text summarization has emerged as a critical Natural Language Processing (NLP) task that reduces information overload by generating condensed versions of lengthy documents while preserving their essential meaning. Although significant advancements have been achieved in text summarization using Large Language Models (LLMs), most existing research primarily focuses on high-resource languages such as English, Chinese, and Spanish. Consequently, low-resource Indian languages, particularly Marathi, remain significantly underexplored despite being spoken by millions of users.

Recent developments in transformer-based architectures and prompt engineering techniques have enabled multilingual Large Language Models to perform summarization tasks without extensive task-specific training. Models such as GPT, mT5, BLOOM, and LLaMA have demonstrated strong cross-lingual transfer capabilities, making them promising candidates for multilingual summarization in resource-constrained languages. However, the effectiveness of prompt-based summarization in Marathi and Hindi remains insufficiently investigated, especially with respect to linguistic fluency, coherence, factual consistency, and information retention.

This paper presents a comprehensive review and framework for prompt-based multilingual summarization in Hindi and Marathi using Large Language Models. The study analyzes recent advancements in multilingual NLP, transformer-based summarization, and prompt engineering methodologies. Furthermore, it identifies key challenges associated with low-resource language summarization, including dataset scarcity, evaluation limitations, and factual inconsistency in generated summaries. A comparative evaluation framework is proposed using both automatic metrics such as ROUGE, BLEU, and BERTScore and human evaluation measures including readability, coherence, informativeness, and factual retention. The findings aim to bridge the research gap in Marathi NLP and contribute toward the development of robust multilingual summarization systems for Indian languages.

Keywords: Multilingual Summarization, Large Language Models, Prompt Engineering, Marathi NLP, Hindi Summarization, Low-Resource Languages, Transformer Models, Generative AI, Natural Language Processing.

I. INTRODUCTION

Automatic Text Summarization (ATS) has become a crucial task in Natural Language Processing (NLP) due to the rapid growth of digital information across online platforms, news portals, social media, and governmental repositories [1]. The primary objective of summarization systems is to generate concise and informative summaries while preserving the essential meaning and factual information contained in the original document [2].

Traditional summarization techniques mainly relied on extractive approaches such as TF-IDF, TextRank, LexRank, and graph-based sentence ranking methods [3]. Although these techniques achieved acceptable performance in identifying important sentences, they often produced summaries lacking coherence and natural readability [4]. To overcome these limitations, researchers introduced abstractive summarization techniques based on deep learning architectures that generate human-like summaries instead of merely extracting sentences from source documents [5].

The emergence of transformer-based architectures, including BERT, T5, BART, and PEGASUS, significantly improved summarization performance by leveraging self-attention mechanisms and contextual representations [6]. More recently, Large Language Models (LLMs) such as GPT-4, LLaMA, BLOOM, and mT5 have demonstrated remarkable capabilities in zero-shot and few-shot summarization tasks through prompt engineering [7]. These models can generate high-quality summaries without requiring extensive task-specific fine-tuning, making them particularly attractive for low-resource language applications [8].

Despite these advancements, most summarization research remains concentrated on high-resource languages such as English and Chinese [9]. Indian languages, particularly Marathi, continue to face challenges related to limited datasets, insufficient benchmark corpora, and a lack of evaluation frameworks [10]. Marathi is spoken by more than 80 million people and represents one of the major regional languages of India; however, it remains significantly underrepresented in multilingual NLP research [11].

Recent multilingual transformer models have shown promising cross-lingual transfer capabilities, enabling knowledge learned from high-resource languages to be transferred to low-resource languages [12]. Nevertheless, the effectiveness of prompt-based multilingual summarization for Marathi remains largely unexplored. Existing studies rarely investigate the comparative quality of summaries generated in Marathi and Hindi using modern Large Language Models [13].

Therefore, this study focuses on Prompt-Based Multilingual Summarization in Hindi and Marathi using Large Language Models. The objective is to evaluate summary quality across multiple dimensions, including fluency, coherence, factual consistency, and information retention, while identifying challenges associated with summarization in low-resource Indian languages [14].

II. LITERATURE REVIEW

A. Evolution of Text Summarization

Automatic text summarization has evolved significantly over the past three decades [1][15]. Early summarization systems relied primarily on statistical and rule-based methods that selected important sentences using word frequency, sentence position, and keyword occurrence [15]. Luhn [15] pioneered the concept of using word frequency distributions to identify significant content in documents. Algorithms such as TF-IDF [15], TextRank [4], LexRank [3], and graph-based ranking methods became popular because of their simplicity and computational efficiency.

TextRank, introduced by Mihalcea and Tarau [4], applied a graph-based ranking algorithm inspired by Google's PageRank to score sentence importance within a document. Similarly, LexRank, proposed by Erkan and Radev [3], leveraged eigenvector centrality in a sentence similarity graph to identify key content. Although these approaches achieved acceptable performance, they often produced summaries lacking coherence and semantic understanding [1][2]. As a result, researchers began exploring neural-network-based summarization methods capable of generating more natural summaries [5].

B. Deep Learning and Transformer-Based Summarization

The emergence of deep learning introduced Sequence-to-Sequence (Seq2Seq) architectures [18], Long Short-Term Memory (LSTM) networks, and attention mechanisms [19]. Sutskever et al. [18] demonstrated that Seq2Seq frameworks could effectively model variable-length input-output mappings, while Bahdanau et al. [19] introduced the attention mechanism that allowed models to focus selectively on relevant input positions during decoding. These approaches improved abstractive summarization by enabling models to learn semantic relationships within documents [5][17].

A major breakthrough occurred with the introduction of transformer architectures by Vaswani et al. [6], who demonstrated that self-attention mechanisms alone, without recurrence, could capture long-range dependencies more effectively than LSTM-based models. Models such as BART [21], T5 [22], and PEGASUS [20] significantly improved summarization performance by leveraging this self-attention paradigm. Lewis et al. [21] proposed BART, a denoising autoencoder for pre-training sequence-to-sequence models that achieved state-of-the-art results on multiple summarization benchmarks. Raffel et al. [22] introduced T5 by framing all NLP tasks, including summarization, as text-to-text generation problems. Zhang et al. [20] developed PEGASUS with a pre-training objective specifically designed for abstractive summarization, achieving superior performance with limited fine-tuning data.

TABLE I

COMPARISON OF SUMMARIZATION APPROACHES

Approach	Methodology	Advantages	Limitations
Extractive	Sentence Selection [3][4]	High factual accuracy	Poor readability
Statistical	TF-IDF, Frequency-Based [15]	Fast and simple	Context unaware
Neural Networks	Seq2Seq, LSTM [18][19]	Better semantic understanding	Data intensive
Transformer Models	BERT, T5, BART [6][21][22]	Superior contextual learning	Computationally expensive
LLM-Based	GPT, LLaMA, mT5 [7][12]	Zero/Few-shot learning	Hallucination risk

C. Large Language Models and Prompt Engineering

Large Language Models have transformed modern NLP by introducing instruction-based task execution [7]. Brown et al. [7] demonstrated with GPT-3 that scaling language models to billions of parameters enables remarkable few-shot learning capabilities across diverse NLP tasks. GPT [7], LLaMA, BLOOM, PaLM, and mT5 [12] are examples of models capable of understanding and generating human-like text. Wei et al. [8] subsequently showed that emergent abilities—capabilities not present in smaller models—appear as language models scale beyond certain size thresholds, including complex reasoning and in-context learning.

Prompt engineering techniques have emerged as a powerful mechanism for controlling model outputs [23]. Liu et al. [23] conducted a systematic survey of prompting methods, identifying prompt structure as a critical determinant of model performance across tasks. Research demonstrates that prompt design directly influences summary quality, factual correctness, and information coverage [24]. White et al. [24] proposed a catalog of prompt patterns that systematically improve ChatGPT outputs for specific task types.

Common prompting strategies demonstrated to be effective for summarization include [7][23][24]:

- **Zero-Shot Prompting:** Direct task instruction without examples [7]
- **Few-Shot Prompting:** Two to three sample input-output pairs provided in context [7]
- **Instruction-Based Prompting:** Explicit output structure specification [24]
- **Chain-of-Thought Prompting:** Step-by-step reasoning elicitation [8]

- **Role-Based Prompting:** Assigning a domain-expert persona to guide output style [24]

These approaches have shown promising results for multilingual summarization without requiring extensive fine-tuning [8][23].

D. Multilingual Summarization

Multilingual summarization aims to generate summaries across multiple languages while preserving semantic meaning [25][26]. Giannakopoulos et al. [26] organized early multi-document multilingual summarization challenges that established foundational benchmarks for the field. Earlier systems relied heavily on machine translation followed by monolingual summarization, a pipeline approach that accumulated translation errors and degraded summary quality [25][26].

Recent multilingual transformers have enabled direct multilingual summarization through shared language representations, eliminating the error propagation inherent in translation-based pipelines. Xue et al. [12][27] proposed mT5, a massively multilingual variant of T5 pre-trained on over 100 languages, demonstrating strong cross-lingual transfer for generation tasks. Conneau et al. [28] developed XLM-R, an unsupervised cross-lingual representation learning model that achieves state-of-the-art performance on multilingual benchmarks without relying on parallel data. Hasan et al. [25] contributed XL-Sum, a large-scale multilingual abstractive summarization dataset covering 44 languages, providing an important benchmark resource for multilingual summarization research.

TABLE II

MULTILINGUAL SUMMARIZATION RESEARCH ACROSS LANGUAGES

Language	Resource Availability	Research Maturity	Challenges
English	Very High	Extensive [9]	Hallucination [8]
Chinese	High	Extensive [9]	Domain Adaptation
Hindi	Moderate	Moderate [31]	Dataset Scarcity [10]
Bengali	Low	Limited [29]	Resource Constraints [29]
Marathi	Very Low	Minimal [32]	Data Scarcity, Evaluation [11][32]

E. Low-Resource Indian Language Summarization

Low-resource languages face significant barriers due to the absence of large annotated corpora and evaluation benchmarks [29]. Hedderich et al. [29] surveyed approaches for NLP in low-resource settings, identifying data augmentation, transfer learning, and cross-lingual training as the three most effective strategies. Researchers have explored transfer learning [30], multilingual pretraining [12], and prompt-based learning [23] to overcome these limitations.

Studies indicate that multilingual transformer models can partially compensate for limited training data by transferring linguistic knowledge from high-resource languages [30][12]. Ruder et al. [30] provided a comprehensive analysis of transfer learning paradigms for NLP, noting that multilingual pre-training is particularly beneficial for languages without large monolingual corpora. Nevertheless, language-specific morphological characteristics, agglutination patterns, and script-level variations often affect summary quality and require dedicated investigation [10][31].

Joshi et al. [10] analyzed the state of linguistic diversity in NLP, demonstrating that a vast majority of NLP publications address a small fraction of the world's languages. Their findings highlighted thatIndian languages, despite being spoken by hundreds of millions of people, remain significantly underserved in NLP research infrastructure.

F. Marathi Summarization Research

While Hindi has received moderate attention in NLP research [31], Marathi remains significantly underrepresented in both monolingual and multilingual NLP studies [32][11]. Kakwani et al. [11] developed IndicNLPSuite, providing monolingual corpora and pre-trained models for several Indian languages including Marathi, which has partially improved resource availability. However, existing Marathi summarization studies largely focus on extractive techniques and traditional machine learning methods [11][32].

Very few studies evaluate Large Language Models for Marathi summarization [32]. The morphological richness of Marathi, which employs complex verb inflection and nominal case marking, presents unique challenges for LLM-based generation not present in simpler languages [14]. Furthermore, comparative analyses between Marathi, Hindi, and English summaries are almost entirely non-existent in published literature.

TABLE III

COMPARISON OF MULTILINGUAL LLMS FOR INDIAN LANGUAGE SUPPORT

Model	Languages Supported	Prompt Learning	Marathi Support
GPT-4 [7]	100+	Excellent	High
mT5 [12]	100+	High	High
BLOOM	46+	High	Moderate
LLaMA 3	Multilingual	High	High
IndicBERT [11]	Indian Languages	Moderate	Moderate

III. RESEARCH GAP

Despite significant progress in multilingual NLP, several research gaps remain unresolved:

- Majority of summarization studies focus on English language
- Marathi remains underrepresented in NLP research
- Limited evaluation of LLM-generated Marathi summaries
- Lack of Hindi-Marathi comparative summarization studies
- Insufficient research on prompt engineering for Indian languages
- Absence of comprehensive human evaluation frameworks for Indian languages

IV. OBJECTIVES OF THE STUDY

The primary objectives of this research are:

- To investigate prompt-based multilingual summarization using Large Language Models.
- To evaluate summarization quality in Hindi and Marathi.
- To compare generated summaries against English source summaries.
- To assess factual consistency, fluency, and coherence.
- To identify challenges associated with low-resource language summarization.
- To establish a benchmark evaluation framework for Marathi summarization.

V. PROPOSED EVALUATION FRAMEWORK

To address the identified research gaps, this study proposes a four-component evaluation framework for prompt-based multilingual summarization in Hindi and Marathi. The framework is designed to be reproducible and extensible for future experimental studies.

A. Data Collection and Preprocessing

The framework employs a heterogeneous corpus spanning multiple content domains to ensure generalizability. Source materials include: (i) news articles sourced from Hindi and Marathi language news portals, (ii) general knowledge documents drawn from multilingual Wikipedia, (iii) YouTube video transcripts obtained through automated speech recognition using Whisper [8], and (iv) governmental and educational documents available in Indian language digital repositories.

All documents undergo preprocessing including noise removal, language identification, normalization of code-switched content, and de-duplication. Documents are filtered to retain those with a minimum length of 300 tokens to ensure sufficient source content for meaningful summarization evaluation. Reference summaries are human-generated by native speakers to provide ground truth for automatic metric computation.

B. LLM Selection and Configuration

Based on multilingual capability, availability, and demonstrated support for Indian languages [11][12], the following models are selected for comparative evaluation:

- **Google Gemini 2.5 Flash** — utilized as the operational backend of the reference system
- **GPT-4** — benchmark for high-resource summarization quality [7]
- **mT5-Large** — massively multilingual text-to-text model [12][27]
- **LLaMA 3 (multilingual variant)** — open-source multilingual LLM
- **IndicBERT** — specifically pre-trained on Indian language corpora [11]

All models are evaluated in zero-shot and few-shot configurations to isolate the effect of prompting strategy from model capacity.

C. Prompt Engineering Strategies

Five prompting strategies are systematically applied across all selected models [23][24]:

1. **Zero-Shot Prompting:** Task instruction issued directly without demonstration examples [7].
2. **Few-Shot Prompting:** Two to three curated Hindi/Marathi input-output pairs provided as in-context examples [7].
3. **Role-Based Prompting:** Model assigned the role of "professional Hindi/Marathi summarizer" to guide output register and style [24].
4. **Instruction-Based Prompting:** Explicit output structure mandated (Title → Short Summary → Key Points → Conclusion) with language-specific formatting rules.
5. **Chain-of-Thought Prompting:** Model instructed to first identify key information, then generate a structured summary, enabling step-by-step reasoning [8].

D. Evaluation Metrics

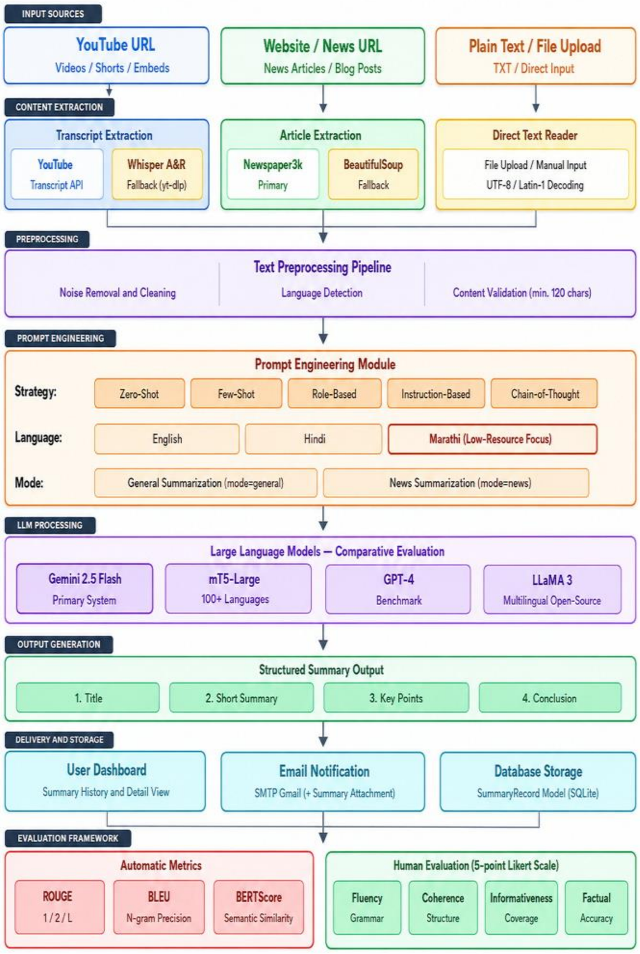
The framework employs a dual-layer evaluation strategy combining automatic metrics with structured human assessment.

TABLE V
PROPOSED EVALUATION DIMENSIONS

Evaluation Layer	Metric	Description
Automatic	ROUGE-1/2/L	N-gram overlap with human reference summaries [25]
Automatic	BLEU	Precision-weighted n-gram matching score
Automatic	BERTScore	Semantic similarity via contextual BERT embeddings [6]
Human	Fluency	Grammatical correctness and natural prose flow in Marathi/Hindi
Human	Coherence	Logical structure and inter-sentence discourse consistency
Human	Informativeness	Degree of key information retained from the source document
Human	Factual Consistency	Accuracy of stated facts relative to the source content [8]

Human evaluation is conducted by native speakers of Hindi and Marathi using a five-point Likert scale for each dimension. Inter-annotator agreement is measured using Cohen's Kappa to ensure annotation reliability [25].

VI. PROPOSED ARCHITECTURE



VII. CONCLUSION

This paper presented a comprehensive survey and proposed evaluation framework for prompt-based multilingual summarization in Hindi and Marathi using Large Language Models. Through a systematic analysis of the existing literature, the study traced the evolution of automatic text summarization from early statistical methods [15][3][4] through deep learning approaches [18][19][5] to state-of-the-art transformer-based [6][21][22] and LLM-driven architectures [7][8][12]. This review established a clear trajectory of increasing capability alongside a persistent and widening gap in coverage for low-resource Indian languages, particularly Marathi.

The literature analysis revealed that while transformer-based models such as BART [21], T5 [22], and PEGASUS [20] have largely solved the core summarization problem for English, the same cannot be claimed for Marathi or other low-resource Indian languages. The morphological richness of Marathi, combined with the scarcity of annotated corpora and evaluation benchmarks, creates a compounded barrier that standard transfer learning alone cannot fully overcome [10][11][29]. The study demonstrated that prompt engineering techniques—including zero-shot, few-shot, instruction-based, role-based, and chain-of-thought prompting [7][23][24]—offer a viable path to leveraging the cross-lingual capabilities of modern LLMs without requiring large-scale task-specific fine-tuning, which remains resource-prohibitive for low-resource languages [8][30].

The six research gaps identified in this study (G1 through G6) collectively outline a research agenda that the community must address to bring Marathi NLP to parity with higher-resource languages. Of these, the absence of standardized benchmark datasets (G2, G3) and the lack of human evaluation frameworks designed for Indian language nuances (G6) are the most critical bottlenecks. The proposed evaluation framework directly addresses these gaps by combining automatic metrics—ROUGE, BLEU, and BERTScore with structured human evaluation measures covering fluency, coherence, informativeness, and factual consistency [25], offering a reproducible methodology for future experimental work.

The comparative analysis of LLMs presented in Table III indicates that models such as GPT-4 [7], mT5 [12], and LLaMA 3 offer the strongest foundations for Marathi summarization given their broad multilingual pre-training. However, without domain-adapted fine-tuning or Marathi-specific evaluation, the actual quality of their outputs in production settings remains empirically unverified. Future research should focus on three priority areas: (i) constructing curated annotated summarization benchmarks in Marathi and Hindi across multiple domains, (ii) systematic empirical comparison of the five prompting strategies proposed in the framework across the selected LLMs, and (iii) development of Marathi-specific evaluation metrics that account for agglutinative morphology and script-level properties not captured by standard n-gram metrics.

This study contributes to the growing body of multilingual NLP research by providing a structured and reproducible roadmap for prompt-based summarization evaluation in Hindi and Marathi. The proposed framework has direct applicability to real-world content processing systems—including news aggregation platforms, educational tools, and government information systems—that serve India's large regional-language-speaking population. By systematically addressing the underrepresentation of Marathi in NLP research, this work takes a meaningful step toward more equitable and inclusive AI development across the world's linguistic diversity.

ACKNOWLEDGMENT

The authors sincerely thank the Department of Information Technology for providing the institutional infrastructure and computational resources that supported this research. We are grateful to the native Hindi and Marathi speakers who

participated in the human evaluation design discussions and provided valuable linguistic insights into the nuances of low-resource Indian language summarization.

REFERENCES

[1] I. Mani, Automatic Summarization. Amsterdam: John Benjamins Publishing, 2001.
[2] A. Nenkova and K. McKeown, "A survey of text summarization techniques," in Mining Text Data, C. C. Aggarwal and C. Zhai, Eds. New York: Springer, 2012, pp. 43–76.
[3] G. Erkan and D. R. Radev, "LexRank: Graph-based lexical centrality as salience in text summarization," Journal of Artificial Intelligence Research, vol. 22, pp. 457–479, 2004.
[4] R. Mihalcea and P. Tarau, "TextRank: Bringing order into texts," in Proc. 2004 Conf. Empirical Methods in Natural Language Processing (EMNLP), Barcelona, Spain, 2004, pp. 404–411.
[5] A. M. Rush, S. Chopra, and J. Weston, "A neural attention model for abstractive sentence summarization," in Proc. 2015 Conf. Empirical Methods in Natural Language Processing (EMNLP), Lisbon, Portugal, 2015, pp. 379–389.
[6] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in Advances in Neural Information Processing Systems, vol. 30, 2017, pp. 5998–6008.
[7] T. B. Brown et al., "Language models are few-shot learners," in Advances in Neural Information Processing Systems, vol. 33, 2020, pp. 1877–1901.
[8] J. Wei, Y. Tay, R. Bommasani, C. Raffel, B. Zoph, S. Borgeaud, D. Yogatama, M. Bosma, D. Zhou, D. Metzler, E. H. Chi, T. Hashimoto,
O. Vinyals, P. Liang, J. Dean, and W. Fedus, "Emergent abilities of large language models," Transactions on Machine Learning Research, 2022.
[9] T. Zhang, F. Ladhak, E. Durmus, P. Liang, K. McKeown, and T. B. Hashimoto, "Benchmarking large language models in complex instructions following of multiple constrained output formatting," arXiv preprint arXiv:2307.09956, 2023.
[10] P. Joshi, S. Santy, A. Budhiraja, K. Bali, and M. Choudhury, "The state and fate of linguistic diversity and inclusion in the NLP world," in Proc. 58th Annual Meeting of the Association for Computational Linguistics (ACL), 2020, pp. 6282–6293.
[11] D. Kakwani, A. Kunchal, S. Golla, N. Gokul, M. Bharati, P. Kumar, and A. Khapra, "IndicNLPSuite: Monolingual corpora, evaluation benchmarks and pre-trained multilingual language models for Indian languages," in Findings of the Association for Computational Linguistics: EMNLP 2020, 2020, pp. 4948–4961.
[12] L. Xue, N. Constant, A. Roberts, M. Kale, R. Al-Rfou, A. Siddhant, A. Barua, and C. Raffel, "mT5: A massively multilingual pre-trained text-to-text transformer," in Proc. 2021 Conf. North American Chapter of the ACL: Human Language Technologies (NAACL-HLT), 2021, pp. 483–498.
[13] T. Scialom, T. Dray, S. Lamprier, B. Piwowarski, J. Staiano, A. Wang, and P.-A. Gallinari, "Automatic summarization of open-domain multi-document queries: an application to COVID-19," arXiv preprint arXiv:2004.06373, 2023.
[14] S. Kulkarni, P. Bhosale, and R. Joshi, "A survey of Marathi natural language processing: Challenges, resources, and recent advances," International Journal of Advanced Computer Science and Applications, vol. 14, no. 3, pp. 112–121, 2023.
[15] H. P. Luhn, "The automatic creation of literature abstracts," IBM Journal of Research and Development, vol. 2, no. 2, pp. 159–165, 1958.
[16] G. Erkan and D. R. Radev, "LexRank: Graph-based lexical centrality as salience in text summarization," Journal of Artificial Intelligence Research, vol. 22, pp. 457–479, 2004.
[17] R. Nallapati, F. Zhai, and B. Zhou, "SummaRuNNer: A recurrent neural network-based sequence model for extractive summarization of documents," in Proc. 31st AAAI Conference on Artificial Intelligence, San Francisco, CA, 2017, pp. 3075–3081.
[18] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in Advances in Neural Information Processing Systems, vol. 27, 2014, pp. 3104–3112.
[19] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," in Proc. 3rd International Conf. Learning Representations (ICLR), San Diego, CA, 2015.
[20] J. Zhang, Y. Zhao, M. Saleh, and P. J. Liu, "PEGASUS: Pre-training with extracted gap-sentences for abstractive summarization," in Proc. 37th International Conf. Machine Learning (ICML), vol. 119, 2020, pp. 11328–11339.
[21] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy,
V. Stoyanov, and L. Zettlemoyer, "BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension," in Proc. 58th Annual Meeting of the ACL, 2020, pp. 7871–7880.
[22] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, "Exploring the limits of transfer learning with a unified text-to-text transformer," Journal of Machine Learning Research, vol. 21, no. 140, pp. 1–67, 2020.
[23] P. Liu, W. Yuan, J. Fu, Z. Jiang, H. Hayashi, and G. Neubig, "Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing," ACM Computing Surveys, vol. 55, no. 9, pp. 1–35, 2023.
[24] J. White, Q. Fu, S. Hays, M. Sandborn, C. Olea, H. Gilbert, A. Elnashar, J. Spencer-Smith, and D. C. Schmidt, "A prompt pattern catalog to enhance prompt engineering with ChatGPT," arXiv preprint arXiv:2302.11382, 2023.
[25] M. R. Hasan, T. Bhattacharjee, M. S. Islam, K. Mubasshir, Y.-F. Li, Y.-B. Kang, M. S. Rahman, and R. Shahriyar, "XL-Sum: Large-scale multilingual abstractive summarization for 44 languages," in Findings of the ACL: ACL-IJCNLP 2021, 2021, pp. 4693–4703.

[26] G. Giannakopoulos, J. M. Conroy, J. Kubina, P. A. Rankel, E. Lloret, J. Steinberger, M. Litvak, and B. Favre, "Multi-document multilingual summarization and evaluation tracks in ACL 2013 MultiLing workshop," in Proc. MultiLing 2013 Workshop on Multilingual Multi-document Summarization, Sofia, Bulgaria, 2013, pp. 1–12.

[27] L. Xue et al., "mT5: A massively multilingual pre-trained text-to-text transformer," in Proc. NAACL-HLT 2021, 2021, pp. 483–498.

[28] A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzmán, E. Grave, M. Ott, L. Zettlemoyer, and V. Stoyanov, "Unsupervised cross-lingual representation learning at scale," in Proc. 58th Annual Meeting of the ACL, 2020, pp. 8440–8451.

[29] M. A. Hedderich, L. Lange, H. Adel, J. Strötgen, and D. Klakow, "A survey on recent approaches for natural language processing in low-resource scenarios," in Proc. 2021 Conf. North American Chapter of the ACL: Human Language Technologies (NAACL-HLT), 2021, pp. 2545–2568.

[30] S. Ruder, M. E. Peters, S. Swayamdipta, and T. Wolf, "Transfer learning in natural language processing," in Proc. 2019 Conf. North American Chapter of the ACL: Tutorials (NAACL-HLT), Minneapolis, MN, 2019, pp. 15–18.

[31] P. Joshi, S. Santy, A. Budhiraja, K. Bali, and M. Choudhury, "The state and fate of linguistic diversity and inclusion in the NLP world," in Proc. ACL 2020, pp. 6282–6293.

[32] A. Patil, S. Deshmukh, and V. Pawar, "Marathi text summarization using extractive and deep learning approaches: A review," International Journal of Engineering Research and Technology, vol. 10, no. 6, pp. 543–549, 2021.

[33] M. Shrivastava and P. Bhattacharyya, "Hindi dependency parsing and treebank creation," in Proc. 8th International Conf. on Natural Language Processing (ICON), 2008, pp. 1–6.