Multilingual & Cross-Lingual Text Summarization of Marathi and English using Transformer Based Models and their Systematic Evaluation

Shlok Deshpande
Dept. of Computer Engineering,
COEP Technological University Pune
Wellesely Road, Pune, 411005,
Maharashtra, India

Vineet Shinde
Dept. of Computer Engineering,
COEP Technological University Pune
Wellesely Road, Pune, 411005,
Maharashtra, India

Siddharth Chaudhari
Dept. of Computer Engineering,
COEP Technological University Pune
Wellesely Road, Pune, 411005,
Maharashtra, India

Yashodhara V. Haribhakta, PhD
Dept. of Computer Engineering,
COEP Technological University Pune
Wellesely Road, Pune, 411005,
Maharashtra, India

ABSTRACT

The proposed Methodology pioneers an approach to multilingual and cross-lingual text summarization, bridging Marathi and English languages through the innovative deployment and specialized optimization of advanced transformer-based models. The research introduces a novel framework designed to navigate and synthesize the linguistic nuances between these two languages, offering a unique contribution to the field of natural language processing. The utilization of Pegasus, T5, and BART is done for English and IndicBART, mT5, and mBART for Marathi summarization, using M2M-100 for translation, to create a synergistic framework that effectively handles the challenges of cross summarization across languages. The core objective is to perform cross-lingual summarization using these models, enhancing their ability to understand and summarize content across Marathi to English & vice-versa. The methodology includes a combination of multiple vast datasets for training and comprehensive evaluation using ROUGE, BLEU, and BERT metrics to assess summarization quality. Additionally, a novel evaluation metric is introduced, which is a combination of concept coverage, semantic similarity and relevance, tailored for assessing multi and cross-lingual summarization quality between English and Marathi. This project not only aims to advance the field of cross-lingual summarization but also seeks to improve accessibility and foster better understanding across linguistic and cultural boundaries.

General Terms

Natural Language Processing, Summarization, Transformers

Keywords

Natural Language Generation, Multi & Cross-Lingual Summarization, Indic Languages

1. INTRODUCTION

In the expanding field of Natural Language Processing (NLP), the research targets the nuanced task of cross-lingual summarization between Marathi and English. This involves the use of sophisticated machine learning models to address linguistic challenges and facilitate clear communication. Summarization, a core NLP task, simplifies text into its most crucial elements, allowing for quicker and easier understanding.

1.0.1 Significance of this Research. This research taps into the relatively under-explored area of Marathi-English summarization, promoting the exchange of information across cultural and linguistic boundaries. It aims to develop and enhance NLP resources such as parallel corpora and language models, specifically for underrepresented languages like Marathi, thereby promoting linguistic equity in technological applications.

1.0.2 Abstractive Summarization. Focusing on Abstractive Summarization, the approach seeks to innovate beyond traditional extractive methods by generating summaries that emulate human paraphrasing capabilities. Employing advanced deep learning models, the methodology aspires to capture the essence of the original text and express it in new, succinct terms, thus making the summaries more informative and readable.

1.1 Architectural Overview

1.1.1 Transformers and Seq2Seq Models. The backbone of modern NLP applications, especially in text summarization, is

formed by Transformers and sequence-to-sequence (seq2seq) models. Both architectures share an encoder-decoder structure:
Transformers: Introduced by Vaswani et al.[1], the Transformer architecture departs from prior models by relying entirely on a self-attention mechanism to weigh the importance of each word in a sentence, irrespective of their positional distances. This allows the model to generate a contextual representation of the text that is rich and nuanced, facilitating more accurate summaries.
Sequence-to-Sequence (Seq2Seq): Traditional seq2seq models transform input text sequences into output sequences through two main components: an encoder that processes the input text and a decoder that generates the output summary[2]. These models are essential for tasks that require a deep understanding of the content, such as machine translation and summarization.
The integration of Transformer and seq2seq technologies has significantly enhanced the capability of NLP systems to handle complex summarization tasks. These models are particularly effective in producing summaries that are not only concise but also rich in context and meaning, adapting well to different languages and text types.

2. LITERATURE REVIEW

2.1 Multilingual and Cross-Lingual Text Summarization

The field of multilingual and cross-lingual text summarization has seen significant advancements in recent years, propelled by the advent of transformer-based models and the availability of new datasets for many low/mid-resource languages. Early attempts at cross-lingual summarization started from the XLSum[3], a multilingual abstractive summarization dataset for 44 languages. It covered many languages for which no public dataset was available. The dataset was evaluated by fine-tuning mT5[4] and showed competitive results. The need for abstractive summarization in low-resource languages rose over extractive due to requirement for shorter summaries and the training of models on NLG tasks such as Headline Generation, Article/Paper Summarization. Then came CrossSum[5], a cross-lingual summarization dataset for 101 languages, which improved the performance and subsequently created m2o (many-to-one) models for 5 languages and one m2m(many-to-many) model for cross-lingual summarization across all language pairs. Research has been done within Cross-Lingual Summarization of Scholarly Documents from English to Ger, Ita, Chi, and Jap languages[6]. The research focused on translation and summarization aspects, and generating extreme TLDR-like summarizations in the mentioned languages and presented fine-tuned models and datasets.

2.2 Translation and Summarization for Indic Languages

The literature on cross-lingual summarization emphasizes the complexity of the task at hand, highlighting the need for advanced models that can navigate the intricacies of multiple languages simultaneously. Significant work has been done in the translation aspect, in terms of identifying linguistic divergence patterns between Marathi and English [7], and various approaches and models have been formulated and tested.

2.2.1 Summarization. Minimal research has been done within the task of summarization for indic languages. Abstractive summarization was done with the use of IndicBART for Hindi, which achieved a score of 0.544 ROUGE F-1 score on the testing dataset [8]. An automated text summarization system was proposed for marathi which achieved 44.48% compression accuracy but fell into extractive summarization [9]. Multiple papers have been published in Indian language summarization using pre-trained sequence-to-sequence and transformer-based models mentioned above. The research was done in Gujarati, Hindi and Telugu, etc languages using open source datasets[10]. For English and Gujarati, PEGASUS and MBart were observed to give the best results respectively, it was noted that the performance of MT5 is superior to IndicBART for Hindi language after both were fine-tuned[11]. A thesis was also done which resulted in a mono-lingual summarization dataset created in Telugu, and the creation of PMIndiaSum[12] a cross-lingual and parallel dataset for summarization for Indian Languages, including evaluation benchmarks and valuable insights for summarization within Indian Languages.

2.3 Evaluation Metrics in Summarization

Assessing the quality of summaries produced by models is essential for determining the performance of summarization algorithms. Metrics like ROUGE and BLEU[13] have been widely used to measure the overlap between the generated summaries and reference summaries. More recently, BERT-based evaluation metrics have been introduced, offering a more nuanced assessment of semantic similarity between the reference text and the generated text[14]. In the context of evaluating cross summarization, the Cross Sum Paper introduced LaSE, an embedding-based metric designed to judge the quality of summaries produced by models [5].

3. RESEARCH GAPS

Despite significant advances in multilingual and cross-lingual text summarization, critical research gaps remain, particularly for under-resourced languages like Marathi.

3.0.1 Research Gaps in Marathi Summarization. While progress has been noted in languages such as English and Hindi, Marathi has not received comparable attention in NLP research. Key challenges include:

—**Limited Pre-trained Models:** There is a noticeable lack of pre-trained transformer models tailored for the NLG task of summarization in Marathi. This limits the development of effective summarization tools that are crucial for non-English languages.
—**Integration Challenges:** Effective integration of translation and summarization models remains elusive, often resulting in a loss of contextual meaning when transferring content between languages. There's a pressing need for frameworks that can seamlessly merge these tasks while preserving the integrity and nuance of the original texts.
—**Customized Summarization Models:** The absence of a robust multilingual summarization model specifically designed for Marathi and other low-resource Indian languages hinders progress. Furthermore, the potential of these models to be trained on large and diverse datasets has yet to be fully explored and evaluated.
—**Inadequate Evaluation Metrics:** Current evaluation metrics predominantly focus on surface-level word matching rather than deep semantic similarity. Particularly in cross-lingual settings, there is a scarcity of metrics tailored for languages like Marathi,

which complicates the assessment of summary relevance and quality.

4. OBJECTIVES

—Implementation and Performance Analysis of Marathi and English Cross & Multi-Lingual and Summarization.
—Performance Analysis of the Proposed Methodology
—Comprehensive Evaluation Metric Development for English & Marathi
—Hindi Language Case Study Implementation

5. METHODOLOGY

To address the identified gaps in cross-lingual text summarization between Marathi and English, the proposed Method includes:-

—Apply dataset for summarization covering a broad range of domains and vocabularies for training, testing, validation, and evaluation.
—Data Cleaning and Analysis to achieve optimised results.
—Generating models on the datasets to be able to perform 2 important tasks:
  —Marathi to Marathi Summarization
  —English to English Summarization
—Generating Models for Monolingual Summarization in Marathi & English
  —English - PEGASUS, T5, BART
  —Marathi - IndicBART, mT5, mBART
—Models for Bilingual Translation from Marathi to English.
  —Helsinki-NLP, M2M-100
—Selection of the models providing the best result based on the evaluation metrics
—Achieving cross lingual summarization through the following approaches:
  —Translating from the source language to target language first and then summarizing[6].
  —Summarizing in the the source language first and then translating from the source language to the target language [6].

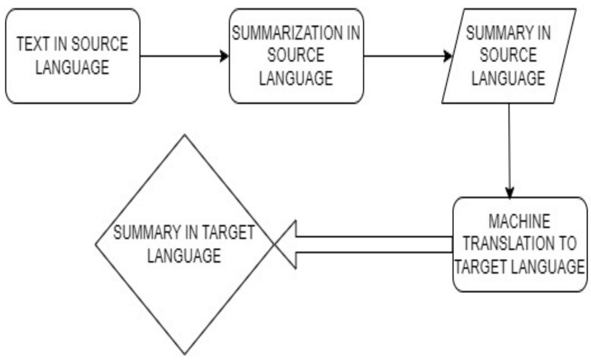

Fig. 1. Approach 1

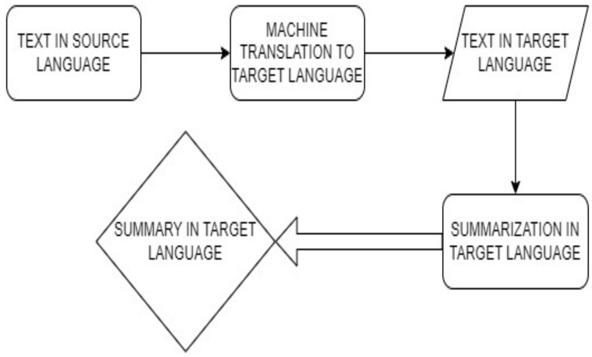

Fig. 2. Approach 2

—Evaluating the approaches and analysing the results on the task of cross lingual summarization.
—Evaluating which combination of translation and summarization models displays the best performance and using that as the standard model for further evaluations.
—Creation of a semantic focused evaluation metric specifically tailored for Marathi and English cross and multilingual summarization, focusing on Concept Coverage, Semantic Similarity and Relevance between the text and summary.
—Evaluation of the reference texts and the generated summaries using the proposed Evaluation Metric.

5.0.1 Models. The models selected for this paper are chosen based on their proven efficacy in NLP tasks related to summarization and translation:

—**English Summarization Models:**
  —**PEGASUS:** Optimized specifically for summarization by pre-training on a masked language modeling objective that mimics a summarization task.
  —**T5 (Text-to-Text Transfer Transformer):** Applies a unified approach to NLP tasks, treating every problem as a text-to-text problem.
  —**BART:** Utilizes a denoising autoencoder strategy by corrupting text with an arbitrary noising function and learning to reconstruct the original text.
—**Marathi Summarization Models:**
  —**IndicBART:** Adapted specifically for Indian languages, it enhances the quality of summaries by understanding nuanced language-specific elements.
  —**mT5:** A multilingual version of T5, capable of understanding and generating text across multiple languages, including Marathi.
  —**mBART:** Pre-trained on large-scale multilingual data, it is well-suited for translation and summarization tasks in multiple languages.
—**Translation Models:**
  —**Helsinki-NLP:** Part of the OPUS project, known for its wide coverage and effective translation across numerous languages.
  —**M2M-100:** Designed to directly translate text between multiple languages without relying on English as a pivot.

5.0.2 Datasets. The Cross-Sum dataset is used, which is specifically curated for the task of cross-lingual summarization. This dataset includes a substantial number of language pairs and article-summary samples, providing a robust foundation for training and evaluating the models.

5.0.3 Evaluation Metrics. The approach integrates traditional metrics with a novel evaluation framework to comprehensively assess model performance:

—**Traditional Metrics:**
  —**ROUGE:** Measures the overlap of n-grams between the generated summaries and a set of reference summaries, assessing the content quality.
  —**BLEU:** Evaluates the precision of n-grams in machine-translated text against reference translations and includes a penalty for overly brief translations.
  —**BERTScore:** These leverage embeddings from pre-trained BERT models to evaluate the semantic similarity between texts.

—**Proposed Cross-Summarization Evaluation Metric:**
  —**Semantic Similarity:** Uses language-agnostic sentence embeddings to compare the semantic content between the source text and its summary.
  —**Relevance:** Applies fine-tuned multilingual model on the downstream task of binary classification using the Cross Sum dataset to assess the relevance of the summary to the original text, ensuring that essential information is preserved.
  —**Concept Coverage:** Analyses the extent to which key concepts and named entities from the original text are covered in the summary, providing a measure of informativeness.

6. EXPERIMENTAL SETUP

6.0.1 Dataset Composition. The Cross Sum dataset encompasses a variety of data splits and sizes, with distinct sections for Marathi, English, and cross-lingual tasks. The dataset consisted of source and target urls along with the summary and text. The average size of the texts in Marathi is 630 words and the summaries are 26 words, while in English the texts are 460 words, and the summaries are of 22 words. The training set was used to custom train the models, the testing set was used to evaluate the models, and the validation set was used to tune the hyper-parameters (Table 1):

Table 1. Dataset Sizes for Summarization

Dataset	Train	Test	Val
Mar Summarization	10,558	1,188	1,254
Eng Summarization	65,000	5,000	5,000
Mar & Eng	1,171	-	-

6.1 Monolingual Summarization Setup

For Marathi to Marathi and English to English summarization, the fine-tuning of transformer-based models is done using a sequence-to-sequence framework. The process includes tokenization, parameter setting, and iterative experimentation to optimize model performance.

6.1.1 Marathi Summarization. Models such as IndicBART, mT5, and mBART were adapted to the Marathi language (Table 2):

6.1.2 English Summarization. English text summarization utilized models like Pegasus, T5, and BART, fine-tuned for optimal summarization capabilities (Table 3):

6.1.3 Translation Models Setup. Translation tasks were performed using M2M-100 and Helsinki-NLP models, applied according to different summarization approaches (Table 4):

6.2 Evaluation Phase

The evaluation of generated summaries will be done using traditional metrics—ROUGE, BLEU, and BERTScore—and the custom Evaluation Metric tailored for abstractive summarization. The generated summaries won't be similar word to word to the reference summaries hence the ROUGE, BLEU evaluation metric scores will be low, and although the BERT scores dictate semantic similarity, they aren't designed for comparing the summaries and the texts. In contrast, the Evaluation Metric emphasizes semantic and contextual coherence, accommodating the shorter lengths and varied phrasing of abstractive summaries compared to original texts. This comprehensive approach helps assess model adaptability and robustness, providing insights into how different summarization models perform under a multilingual setup, thereby enhancing the understanding of their effectiveness across varied linguistic contexts.

6.3 Evaluation Metric Experimentation Setup

To fine-tune the MuRIL model for the downstream task of binary classification for relevance, the CrossSum dataset was used, encompassing both multilingual and cross-lingual sections. This dataset was meticulously prepared by marking relevant summary-text pairs with a relevance score of 1 and creating non-relevant pairs (scored as 0) by shuffling texts and re-pairing them with the original summaries. Additionally, to construct the cross-lingual dataset, English and Marathi texts and summaries were merged, indicating non-relevance, while using the English-Marathi CrossSum dataset for relevant pairs. The merged dataset comprised 35,242 relevant and 55,242 irrelevant summaries, designed to maintain computational efficiency and balance (Table 5).

6.3.1 Model Selection & Conceptualization. For model selection and conceptualization, the integration of language-specific Named Entity Recognition (NER) models was done to ensure the preservation of key factual content in summaries. The English summaries were analysed using the "dbmdz/bert-large-cased-fine-tuned-conll03-english" model[15], and the Marathi summaries with the "l3cube-pune/marathi-ner" model[16]. These models are crucial for identifying essential entities like persons, locations, and organizations, thus maintaining the integrity of the source texts in the summaries.
MuRIL[17] was chosen for relevance calculations due to its robust performance in multilingual binary classification tasks(Table 6). This model is particularly adept at handling the linguistic nuances of Indian languages, making it ideal for assessing the alignment between the content of summaries and the narratives or arguments in the full texts. Additionally, the "setu4993/LaBSE" model was used for its efficacy in generating language-agnostic sentence embeddings, allowing us to compare the semantic content of the original texts and their summaries.
The evaluation metric was finely calibrated through rigorous experimentation, merging various components into a cohesive pipeline shown in Figure 3.

Table 2. Marathi Model Training Parameters

Model Name	Epochs	Batch Size	Learning Rate	Model Size	Max Sequence Length
mBART	4	4	2e-5	611M	1024
IndicBART	8	8	0.001	244M	512
mT5	8	8	5.6e-5	300M	512

Table 3. English Model Training Parameters

Model Name	Epochs	Batch Size	Learning Rate	Max Sequence Length
Pegasus	4	8	5.6e-5	512
BART	4	8	5.6e-5	512
T5	8	8	5.6e-5	512

Table 4. Translation Models

Model Name	Approach 1	Approach 2
M2M-100	m2m100-1.2B	m2m100-418M
Helsinki-NLP	MR to EN	EN to MR

Table 5. Relevance Dataset

Criteria	Train	Validation
Dataset Size	81,436	9,048

Table 6. Relevance Model Training Parameters

Model	Epochs	Batch Size	Learning Rate
MuRIL	10	64	5.6e-5

7. RESULTS AND DISCUSSION

7.1 Marathi to Marathi Summarization

Table 7 demonstrates significant improvements in Marathi text summarization post-model fine-tuning, especially with mbart-large-50 showing substantial gains across all metrics. These results highlight the challenges and complexity of Marathi summarization due to its intricate linguistic features, especially when the models have limited contextual data to learn from.

7.2 English Summarization Results

Pegasus outperformed other models in English summarization, achieving the highest scores across all metrics with an R-1 of 46.89 and a BERTScore of 0.81. This reflects its robust capability across comprehensive linguistic metrics. The results, as shown in Table 8, indicate the effectiveness of these models in managing the complexities of the English language.

7.3 Evaluation Metric Results

The MuRIL model exhibited high accuracy, with 99.56% on validation and 99.24% on unseen datasets, demonstrating its effectiveness in relevance classification between English and Marathi texts. The Model Training Graph is shown in Figure 5. The weights in the metric after experimentation were decided as follows w1(similarity score) 0.3, w2(concept coverage) 0.2 and w3(relevance) 0.5. The evaluation metric was tested on a snippet of the CrossSum English Summarization dataset where it gave a score of 0.88 and for the Marathi Summarization it gave a score of 0.86. Examples of the summaries generated from the peagsus and m-bart fine-tuned models are given in figure 4 reinforcing the need for the proposed evaluation metric.

7.4 Evaluation of Cross-Lingual Summarization Approaches

Table 9 shows that approaches using the M2M-100 translation model significantly outperform those using Helsinki, primarily due to M2M-100's larger size. Approach-1 consistently achieves better results in cross-lingual summarization between English and Marathi, as it more accurately translates shorter texts with fewer errors. Despite this, the scores from the evaluation metrics range from 0.70 to 0.80, limited by M2M-100's tendency to use simpler language, which affects the contextual richness of the summaries.

7.5 Hindi Language Case Study

A case study was conducted on the m-bart fine-tuned model for Marathi to evaluate its performance on Hindi texts, given the linguistic similarities between Hindi and Marathi. The experiment utilized the Cross Sum Hindi summarization dataset, which consisted of 31240 training, and 2370 testing and validation samples.
Initially, the model processed Hindi texts and output summaries in Marathi, showing some contextual alignment but poor grammatical quality. Despite fine-tuning adjustments (2 epochs, learning rate 5.6e-5), improvements in validation loss did not enhance performance. The model struggled to differentiate between Hindi and Marathi due to their semantic similarities, suggesting a need for dedicated fine-tuning on Hindi data from scratch to achieve accurate Hindi summarization.

7.5.1 Challenges Faced. The research encountered significant challenges, including model overfitting, translation inaccuracies, and the complexities of abstractive summarization evaluation. Delays in API access and the limitations of smaller GPUs also hindered progress, requiring adjustments in the approach to fine-tuning and evaluation.

8. CONCLUSION

8.1 Cross & Multi-Lingual Summarization

This study has advanced NLP for Marathi and English by leveraging top datasets to fine-tune summarization models, evaluated via ROUGE, BLEU, and BERT scores. PEGASUS and M-BART emerged as superior models for English and Marathi, respectively. These models, paired with translation models M2M-100 and Helsinki, showed that the first approach of summarizing then translating is more effective, especially with M2M-100 due to fewer translation errors. A Hindi case study revealed that semantic similarities led to poor grammatical outputs, suggesting the need for isolated fine-tuning.

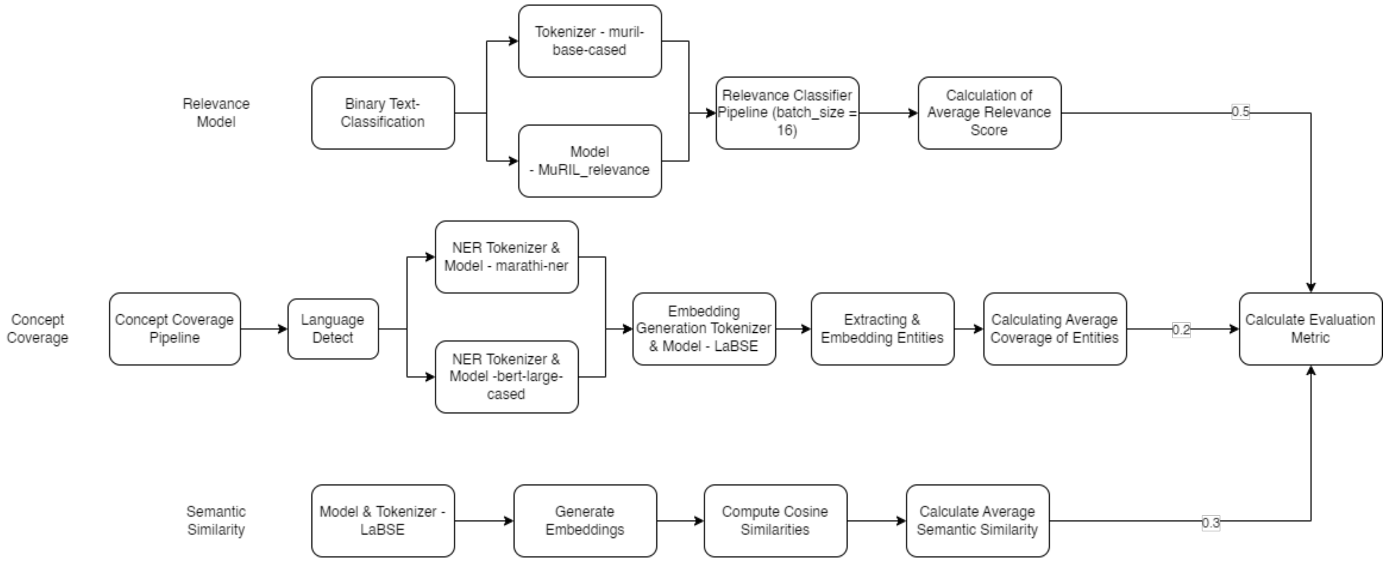

Fig. 3. Evaluation Metric Architecture

Table 7. Marathi Summarization Results

Model Name	R-1	R-2	R-L	BLEU	BERT
mT5-small	-	-	-	-	-
mT5-small(fine-tuned)	5.3	0.54	5.3	10	0.75
IndicBART	4.61	0.73	4.58	0.78	0.6567
IndicBART(fine-tuned)	6.3	0.82	6.3	12.5	0.75
mbart-large-50	5.92	0.87	5.9	7.89	0.72
mbart-large-50(fine-tuned)	22.82	7.8	22.82	27.2	0.81

Languages	Reference Summaries	Generated Summaries
English	Russian President Vladimir Putin has sought to allay Israeli concerns at Russia's military build-up in Syria.	Russian President Vladimir Putin has defended his country's military presence in Syria, amid reports that it has sent hundreds of troops to the country.
Marathi	काँग्रेस-राष्ट्रवादी काँग्रेस आणि शिवसेनेसारखे संपूर्ण वेगळ्या विचारधारांचे पक्ष सरकार स्थापन करत असल्याने एका बाजूला या सरकारच्या भवितव्याबाबत साशंकता व्यक्त केली जात आहे, तर दुसरीकडे सर्वसामान्य नागरिक नवीन सरकारकडून अनेक अपेक्षा व्यक्त करत आहेत	शिवसेना, काँग्रेस आणि राष्ट्रवादी काँग्रेस या तिन्ही पक्षांच्या जाहीरनाम्यात शेतकऱ्यांना कर्जमाफीची घोषणा करण्यात आली होती या घोषणाबाजीनंतर नवीन सरकारकडून अपेक्षा व्यक्त करण्यात आल्या आहेत

Fig. 4. Example Summaries Output

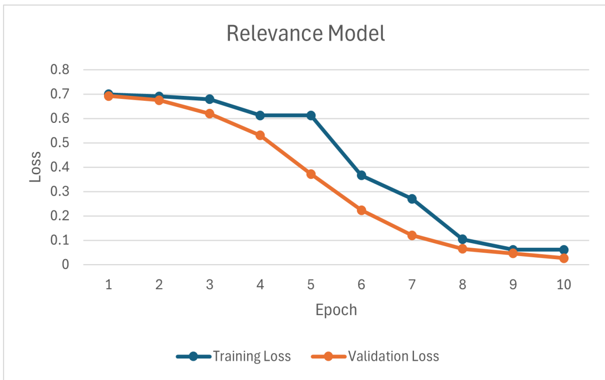

Fig. 5. Model Training Graph

8.2 Evaluation Metric

Recognizing the limitations of traditional metrics in assessing abstractive summarization, a new Evaluation Metric focusing on Relevance, Concept Coverage, and Semantic Similarity was developed. This metric, validated with high accuracy by the MuRIL model, provided a multi-dimensional assessment of summaries, crucial for enhancing cross-lingual systems. It ensures that summaries are contextually relevant and reliable, accommodating the possibility of numerous acceptable interpretations for extensive texts.

8.3 Future Work

Future research will address current limitations, such as the inadequate performance of available translation models and the need for more comprehensive datasets. Improvements in these areas are expected to enhance Cross-Lingual Summarization outcomes significantly. Additionally, developing more intricate evaluation metrics to assess summaries' grammatical and contextual quality will push the boundaries of research in Cross-Lingual Summarization, particularly for underrepresented languages like Marathi. This ongoing work aims to refine the tools

Table 8. English Summarization Results

Model Name	R-1	R-2	R-L	BLEU	BERT
Pegasus	46.89	23.72	39.13	17.21	0.81
BART	44.07	21.01	35.51	14.06	0.79
T5	37.93	15.43	30.08	9.6	0.77

Table 9. Cross-Lingual Summarization results

Translation	**Approach**	**Translation Model**	**Summarization Model**	**Score**
En Text to Mr Summarization	Approach 1	M2M-100	PEGASUS	0.76
	Approach 2	M2M-100	M-BART	0.72
Mr Text to En Summarization	Approach 1	M2M-100	M-BART	0.73
	Approach 2	M2M-100	PEGASUS	0.71
En Text to Mr Summarization	Approach 1	HELSINKI	PEGASUS	0.25
	Approach 2	HELSINKI	M-BART	0.20
Mr Text to En Summarization	Approach 1	HELSINKI	M-BART	0.23
	Approach 2	HELSINKI	PEGASUS	0.18

and methods necessary for overcoming the remaining challenges in this evolving field.

9. REFERENCES

[1] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. *CoRR*, abs/1706.03762, 2017.

[2] Ilya Sutskever, Oriol Vinyals, and Quoc V. Le. Sequence to sequence learning with neural networks, 2014.

[3] Tahmid Hasan, Abhik Bhattacharjee, Md. Saiful Islam, Kazi Mubasshir, Yuan-Fang Li, Yong-Bin Kang, M. Sohel Rahman, and Rifat Shahriyar. XL-sum: Large-scale multilingual abstractive summarization for 44 languages. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 4693–4703, Online, August 2021. Association for Computational Linguistics.

[4] Linting Xue, Noah Constant, Adam Roberts, Mihir Kale, Rami Al-Rfou, Aditya Siddhant, Aditya Barua, and Colin Raffel. mt5: A massively multilingual pre-trained text-to-text transformer. *CoRR*, abs/2010.11934, 2020.

[5] Abhik Bhattacharjee, Tahmid Hasan, Wasi Uddin Ahmad, Yuan-Fang Li, Yong-Bin Kang, and Rifat Shahriyar. CrossSum: Beyond English-centric cross-lingual summarization for 1,500+ language pairs. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2541–2564, Toronto, Canada, July 2023. Association for Computational Linguistics.

[6] Sotaro Takeshita, Tommaso Green, Niklas Friedrich, Kai Eckert, and Simone Paolo Ponzetto. X-scitldr: Cross-lingual extreme summarization of scholarly documents. In *2022 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, pages 1–12, 2022.

[7] S. Kulkarni, Prapti Deshmukh, Majharoddin Kazi, and Karbhari Kale. Linguistic divergence patterns in english to marathi translation. *International Journal of Computer Applications*, 87, 01 2014.

[8] Arjit Agarwal, Soham Naik, and Sheetal S. Sonawane. Abstractive text summarization for hindi language using indicbart. In *Fire*, 2022.

[9] VAISHALI P KADAM1 SAMAH ALI ALAZANI and C NAMRATA MAHENDER. A text summarization system for marathi language.

[10] Ashok Urlana. *Enhancing Text Summarization for Indian Languages: Mono, Multi and Cross-lingual Approaches*. PhD thesis, 07 2023.

[11] Ashok Urlana, Sahil Manoj Bhatt, Nirmal Surange, and Manish Shrivastava. Indian language summarization using pretrained sequence-to-sequence models, 2023.

[12] Ashok Urlana, Pinzhen Chen, Zheng Zhao, Shay B. Cohen, Manish Shrivastava, and Barry Haddow. Pmindiasum: Multilingual and cross-lingual headline summarization for languages in india, 2023.

[13] Kishore Papineni, Salim Roukos, Todd Ward, and Wei Jing Zhu. Bleu: a method for automatic evaluation of machine translation. 10 2002.

[14] Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q. Weinberger, and Yoav Artzi. Bertscore: Evaluating text generation with BERT. *CoRR*, abs/1904.09675, 2019.

[15] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding, 2019.

[16] Parth Patil, Aparna Ranade, Maithili Sabane, Onkar Litake, and Raviraj Joshi. L3cube-mahaner: A marathi named entity recognition dataset and bert models, 2022.

[17] Simran Khanuja, Diksha Bansal, Sarvesh Mehtani, Savya Khosla, Atreyee Dey, Balaji Gopalan, Dilip Kumar Margam, Pooja Aggarwal, Rajiv Teja Nagipogu, Shachi Dave, Shruti Gupta, Subhash Chandra Bose Gali, Vish Subramanian, and Partha Talukdar. Muril: Multilingual representations for indian languages, 2021.