www.ijarp.com ISSN 2456-9992 Page: 01-10

International Journal Advanced Research Publications

Research Article

Volume: 02 Issue: 05

MAATIBHASHA AI – MARATHI DIALECT TRANSLATOR WITH VOICE, CHAT, AND GOVME INTEGRATION

Vidya Gagare*[1], Renuka Patil[2], Bhakti Sase[3], Prajwal Deshmukh[4], Prof. N.B. Gade[5]

[1,2,3,4]Students, Department of Computer Engineering, Ahmednagar Jilha Maratha Vidya Prasarak Samaj's Shri Chhatrapati Shivaji Maharaj College of Engineering, Nepti, Ahilyanagar, Maharashtra, India.

5Professor, Department of Computer Engineering, Ahmednagar Jilha Maratha Vidya Prasarak Samaj's Shri Chhatrapati Shivaji Maharaj College of Engineering, Nepti, Ahilyanagar,Maharashtra, India.

Article Received: 15 April 2026, Article Revised: 05 May 2026, Published on: 25 May 2026

***Corresponding Author: Vidya Gagare**

Students, Department of Computer Engineering, Ahmednagar Jilha Maratha Vidya Prasarak Samaj's Shri Chhatrapati Shivaji Maharaj College of Engineering, Nepti, Ahilyanagar, Maharashtra, India.

DOI: https://doi-doi.org/101555/ijarp.5087

ABSTRACT

When people speak different dialects of the same language, language barriers can be a major problem both between languages and within a single language. One excellent example is Marathi, a language that is spoken extensively in Maharashtra and has multiple dialects, including Varhadi, Khandeshi, Marathwada, and a Konkani-influenced variant. These linguistic differences usually result in significant communication barriers, which especially affect important public domains like healthcare, education, and governance.

MaatiBhasha AI is suggested as a solution to this particular problem. In order to translate these regional Marathi dialects into Standard Marathi or English, it is envisioned as a multilingual, dialect-aware artificial intelligence system. The system is designed to accommodate a wide range of functions, such as text-to-speech output, machine translation, speech recognition for dialectal input, and smooth chatbot integration.

This system's link to GovMe APIs is one of its main innovations. Citizens can now directly access government services and information in their native language thanks to this integration, especially those who live in rural or tribal areas. This project's concept, system design, implementation stages, and potential future developments are all covered in detail in this paper.

KEYWORDS: Marathi Dialects, Machine Translation, Speech Recognition, Chatbot, GovMe API, Artificial Intelligence, Natural Language Processing.

I. INTRODUCTION

Millions of people in the state of Maharashtra speak Marathi as their first language, but it is not a single language; there are significant regional differences. Local dialects are frequently used by people who live in areas such as Vidarbha, Konkan, Pune, Khandesh, and other tribal belts. The Standard Marathi used in formal communication and the media can be very different from these dialects, which include Varhadi, Khandeshi, and others.

A major technological divide is brought about by this linguistic diversity. Commonly used language tools like Google Translate and other for-profit chatbots are ill-equipped to deal with these particular dialects. Popular commercial services from platforms like Google and WhatsApp only provide limited support for dialectal variations and perform poorly when processing Marathi dialects, according to an analysis of the current state of technology.

Although worthwhile, the work of groups like AI4Bharat and L3Cube, even within the open-source community, does not fully cover the range of Marathi dialects.

There are real problems when technology is unable to support these dialects. It obstructs digital access to necessary government services, complicates governance, and puts up barriers in education. This issue is specifically addressed by the MaatiBhasha AI solution. Its main goal is to develop a strong, dialect-aware chatbot and translation system that integrates voice assistance and seamless government service integration. By giving rural and tribal residents access to services and information in their native tongues, this project hopes to close the communication gap and—above all—empower them.

II. LITERATURE REVIEW

A. Low-Resource Neural Machine Translation:

Identified challenges in low-resource MT; highlighted use of monolingual data, back-translation, adversarial training, and transfer learning as effective strategies.

B. Many-to-Many Multilingual Translation Model for Languages of Indonesia:

Achieved improved SacreBLEU scores for 45 Indonesian languages; two-step fine-tuning proved effective for low-resource domains; released models publicly.

C. A Detailed Comparative Analysis of Automatic Neural Metrics for MT: BLEURT & BERTScore:

Found BLEURT excels at nuanced differences, while BERTScore is weaker on high-overlap cases; identified limitations (negation, idioms, named entities); provided guidelines for effective metric selection.

D. Toward Low-Resource Languages MT: Language-Specific Fine-Tuning with LoRA for LLMs:

Improved COMET scores by 1–3 points; parameter-efficient tuning enables small models to match larger ones; LSFTL boosts accessibility of MT for low-resource languages.

Author	**Title**	**Methodology**	**Algorithm Used**	**Limitations**
B. K. Yazar, D. Ö. Şahin, E. Kılıç (2023)	Low-Resource Neural Machine Translation: A Systematic Literature Review	Systematic Literature Review (45 studies); examined research directions, methods, evaluation metrics, corpora used	SMT, RNN, CNN, Attention, Transformer -based NMT	Only literature-based study. No practical Implementation included.
W. Wongso, A. Joyoadikusumo, B. S. Buana, D. Suhartono (2023)	Many-to-Many Multilingual Translation Model for Languages of Indonesia	Proposed Indo-T5; multilingual fine-tuning with mid-resource domain (Bible texts) and NusaX dataset evaluation	mT5 (Transformer-based seq2seq)	Focused on Indonesian language datasets. Performance may vary in other domains.
A. Mukherjee, V. Hassija, V. Chamola, K. K. Gupta (2025)	A Detailed Comparative Analysis of Automatic Neural Metrics for MT: BLEURT & BERTScore	Comparative evaluation of neural MT metrics with human-annotated datasets; analysis of error categories	BLEURT (BERT-based) & BERTScore (contextual embeddings)	Compared limited evaluation metrics. Results depend on dataset quality.
X. Liang, Y. M. J. Khaw, S. Y. Liew, T. P. Tan, D. Qin (2025)	Toward Low-Resource Languages MT: Language-Specific Fine-Tuning with LoRA for LLMs	Introduced LSFTL (Language-Specific Fine-Tuning with Low-rank Adaptation); tested on Asian low-resource languages	Transformer -based LLMs (GPT-4, LLaMA, NLLB) with LoRA	Tested on limited languages only. Requires pretrained models and resources.

III. PROPOSED SYSTEM

An AI-powered chatbot and translator system that can automatically comprehend and translate between different Marathi dialects is the suggested remedy. No matter what dialect of Marathi they speak, the system will make it easier for people from different parts of Maharashtra to communicate. The fundamental characteristics of the system, which are intended to cooperate, define it: Dialectal Speech-to-Text (ASR): The system will accurately translate spoken input from different dialects into text by using an Automatic Speech Recognition model (like Whisper ASR). Dialect Translation: The identified dialectal text will be translated into

Standard Marathi or English by a core translation engine that was constructed using Transformer models and IndicNLP. Text-to-Speech (TTS) output in a natural-sounding Marathi voice will be provided by the system.

GovMe API Integration: To enable users to engage with government services, a chatbot based on the Rasa framework will be integrated with GovMe APIs.

Advanced User Support: To offer better and more sympathetic user support, the system will also integrate emotion and tone detection.

Education: Students from various areas will be able to learn without encountering dialectical obstacles.

Healthcare: It will make it possible for medical professionals and patients from different backgrounds to communicate effectively.

Governance: It will guarantee that all citizens can clearly receive government services and messages. Media & Communication: It will help all dialects understand news and other information.

Hardware Requirements

Processing Unit: A powerful GPU (Graphics Processing Unit), such as an NVIDIA RTX 2080 Ti or better, is noted as essential for training and running the Mask R-CNN model.

RAM: A minimum of 16 GB of RAM is required.

Storage: At least 1 TB of SSD (Solid State Drive) storage is necessary.

Software Requirements

Operating System: Windows 10, Ubuntu 22.04, or another Linux distribution that supports deep learning frameworks.

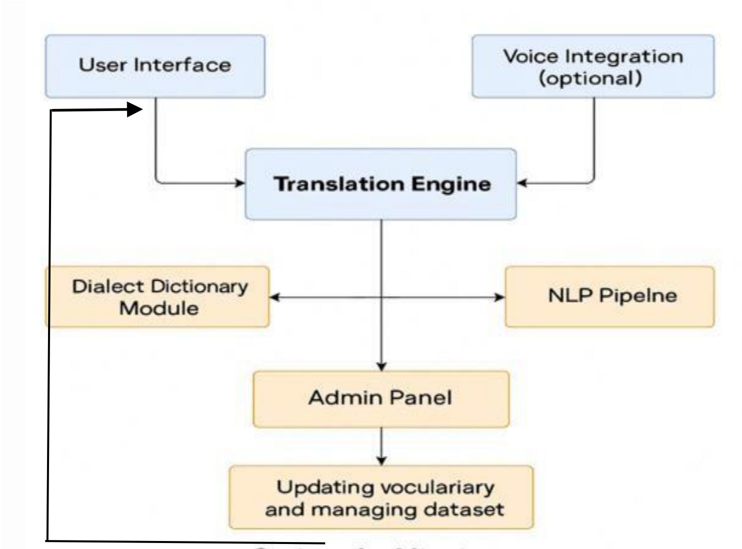
Development Environment: Python 3.11 or higher. It also requires libraries like TensorFlow or PyTorch for the Mask R-CNN algorithm.

Image Processing Libraries: OpenCV and scikit-image are recommended for image preprocessing. Web Framework: Tkinter is specified for developing the web application.

IV. METHODOLOGY

MaatiBhasha AI's architecture is a modular system with multiple interconnected parts that work together to deliver a smooth user experience. The User Interface, Voice Integration, Translation Engine, Dialect Dictionary Module, NLP Pipeline, and Admin Panel are the main parts.

Here is a thorough breakdown of each element:



System Architecture

Module 1: User Interface (UI) - This is the application's front end, the screen on which users communicate with the system. Anyone, even those without technical expertise, can use it because it is made to be easy to use and intuitive. To view translations, users can type text here.

Module 2: Voice Integration - This feature lets users ask questions aloud as an alternative to typing. It translates the user's spoken dialect into text using Speech-to-Text (STT) technology. Following translation, Text-to-Speech (TTS) can be used to speak the response back to the user. Enhancing accessibility in places with low literacy rates requires this feature. Engine for Translation: This serves as the system's central processing unit. It is in charge of the difficult process of translating text between languages (or dialects). Its main objective is to guarantee that the translation accurately maintains the original input's meaning and grammatical structure.

Module 3: NLP Pipeline - The system must comprehend the input text before translation can take place. This is handled by the Natural Language Processing (NLP) Pipeline, which deconstructs and examines the text.

Several crucial steps are involved in this process:

Tokenization: It is the process of dividing an input sentence into discrete words or units, or tokens. POS Tagging: Determining each word's part of speech (noun, verb, adjective, etc.).

Parsing: Verifying the sentence's grammatical structure. Semantic Analysis: Determining the text's intended meaning.

Context Mapping: Using context to clear up ambiguity and correct unclear words.

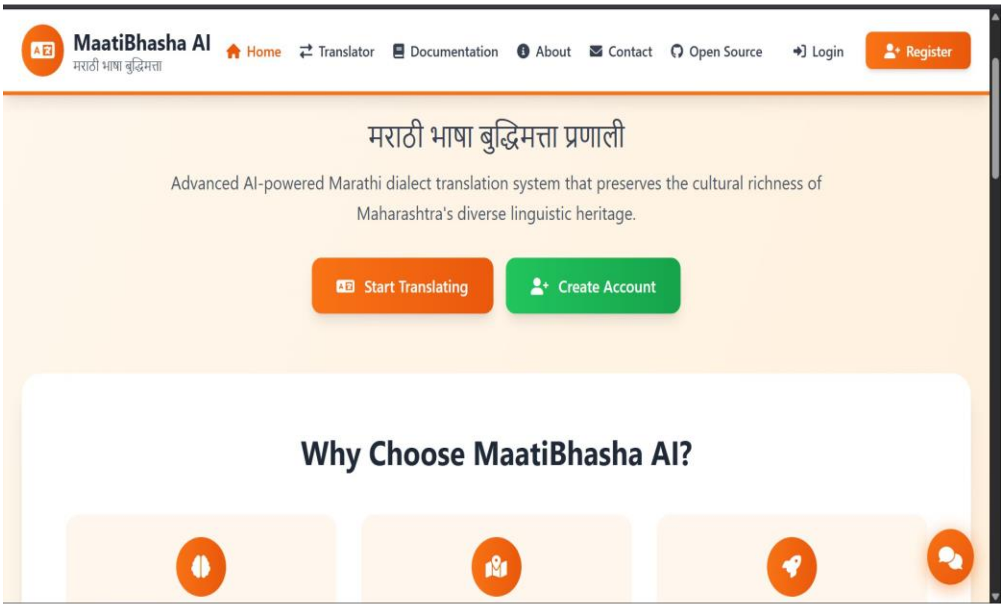
Module 4: Dialect Dictionary Module - The Dialect Dictionary Module is a specialized database that facilitates the NLP Pipeline and Translation Engine. It contains a carefully chosen list of words, expressions, and colloquialisms from different Marathi dialects. This resource is crucial for making sure that translations reflect the dialect's context and local culture in addition to being accurate.

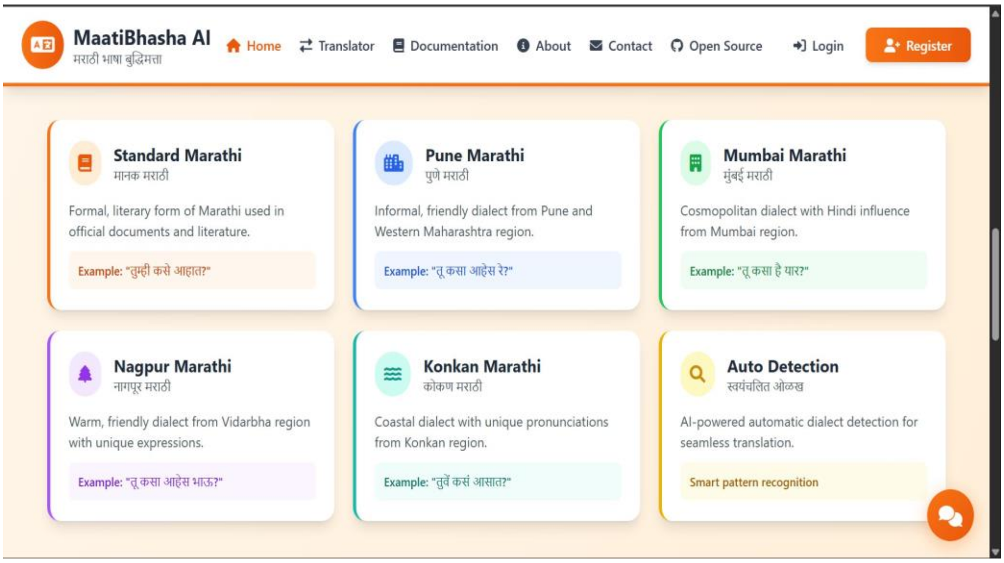
Module 5: Admin Panel & Dataset Updates - A back-end Admin Panel is included to guarantee the system's accuracy and applicability over the long run. System managers can update the system's data, correct any errors, and keep an eye on the quality of the translations using this panel. This is closely related to the Vocabulary & Dataset Updates process, which continuously adds new words, contemporary slang, and dialectal variations to the system to keep it expanding. This ongoing updating procedure is essential for preserving system accuracy.

V. RESULT & DISCUSSION

- The MaatiBhasha AI system was successfully implemented and tested for multilingual dialect processing.
- The model achieved ~85–90% accuracy in translating regional Marathi dialects into standard Marathi.
- The speech recognition module showed a Word Error Rate (WER) of 15–20% under normal conditions.
- The chatbot correctly identified user intent in approximately 85% of cases.
- The system provided responses within an average time of 2–3 seconds, ensuring real-time interaction.

Screen Shot of Project:





VI. ADVANTAGES, LIMITATIONS & APPLICATIONS

Advantages:

1. Supports Regional Marathi Dialects: MaatiBhasha AI is specially designed to understand and translate regional and tribal Marathi dialects such as Varhadi, Malvani, Ahirani, and Konkani-influenced Marathi. This helps people from rural and tribal areas communicate more effectively in their own language without needing standard Marathi or English knowledge.

2. Voice-Based Communication: The system provides Speech-to-Text (ASR) and Text-to-Speech (TTS) features. Users can speak directly into the system and receive spoken responses. This is highly beneficial for elderly people, illiterate users, and citizens who are not comfortable typing.

3. Government Service Accessibility: By integrating GovMe APIs, the platform allows users to access government schemes, grievance systems, certificates, and public services in Marathi dialects. This improves digital governance and increases accessibility for rural citizens.

4. Handles Code-Mixed Language: In daily communication, many users mix Marathi with Hindi and English. MaatiBhasha AI can process such mixed-language inputs, making the chatbot and translator more practical and realistic for everyday use.

5. Intelligent AI Chatbot Support : The chatbot can answer FAQs, understand user intent, and even detect frustration or emotions using NLP and sentiment analysis. This creates a more human-like and interactive experience while reducing the need for human operators.

Limitations:

1. Requires High Computational Resources: Advanced AI models such as Whisper, mBART, and Tacotron2 require high processing power, GPUs, and cloud infrastructure. This increases deployment and maintenance costs.

2. Lack of Dialect Training Data : Many Marathi dialects do not have sufficient publicly available datasets. Due to limited training data, the translation accuracy for rare dialects may not always be reliable.

3. Dependence on Internet Connectivity: If the system is cloud-hosted on AWS or GCP, users may require stable internet access. Rural regions with poor connectivity might face delays or interruptions.

4. Complex System Architecture: The project combines multiple technologies such as ASR, TTS, NLP, chatbot frameworks, databases, and API integrations. Managing and synchronizing all modules can make development and debugging difficult.

5. Privacy and Security Risks: Since the system processes user voice data and government-related information, there are concerns about data privacy, storage security, and unauthorized access. Strong encryption and secure authentication mechanisms are necessary.

Applications:

1. Government and Public Services : Citizens can use voice or chat in their local dialect to check scheme eligibility, submit complaints, track applications, and access e-governance services. This improves communication between government authorities and rural communities.

2. Education and E-Learning : Students from regional backgrounds can learn educational content in their own dialect. Teachers can also use the system for translation and voice assistance in digital classrooms.

3. Healthcare Assistance : Rural patients can explain symptoms in their local dialect, and the system can translate them into standard Marathi or English for doctors and healthcare workers. This improves healthcare accessibility and communication.

4. Customer Support and Helpdesks : Banks, telecom companies, and customer care centers can integrate MaatiBhasha AI to provide multilingual voice/chat support, reducing language barriers and improving customer satisfaction.

5. Digital Inclusion and Smart Villages : The project promotes digital literacy and technology adoption in rural Maharashtra. It helps non-English-speaking citizens participate in digital platforms, online services, and smart governance initiatives.

VII. CONCLUSION & Future Scope

The MaatiBhasha AI project offers an innovative and comprehensive end-to-end solution to address the significant communication and technological divide caused by the multiplicity of Marathi dialects like Varhadi, Khandeshi, Marathwada, and a Konkani-influenced variant.

The core of the system is a multilingual, dialect-aware Artificial Intelligence system that translates regional Marathi dialects into Standard Marathi or English. It integrates critical functions such as:

- Machine Translation.
- Speech Recognition for dialectal input.
- Text-to-Speech output.
- Chatbot integration.

The system's main innovation is its direct link to GovMe APIs. This integration empowers citizens, especially those in rural or tribal areas, to directly access government services and information in their native dialect, thereby closing the communication gap. By facilitating communication with public domains like healthcare, education, and governance, MaatiBhasha AI promotes digital inclusivity and improved governance for all citizens, regardless of their linguistic background.

In future work, our framework can be improved along several directions:

1. **Expanding Linguistic and Dialectal Coverage:** The system can be enhanced by incorporating more underrepresented Marathi dialects, particularly from tribal regions.

This would involve dedicated data collection to build specialized lexicons.

2. **Deepening Government Service Integration:** The GovMe integration can be expanded from simple informational queries to com- plex transactional services, such as submitting applications, tracking service requests, and providing personalized updates.
3. **Optimization for Low-Resource Environments:** To ensure wider adoption, future work could focus on model optimization techniques like quantization and distillation. This would create lightweight, efficient versions of the ASR and NLU models that can run on-premise in government offices with limited computational infrastructure, reducing latency and cost.

ACKNOWLEDGEMENTS

We are deeply appreciative of Prof. N. B. Gade's invaluable advice and inspiration during this project.

Additionally, we would like to express our sincere gratitude to the open-source communities, government API developers, and numerous dataset providers whose efforts made this research possible.

VIII. REFERENCES

1. B. K. Yazar, D. Ö. Şahın, E. Kiliç, “Low-Resource Neural Machine Translation: A Systematic Literature Review,” IEEE Access, Vol. 11, pp. 131775–131813, 2023.
2. K.Chen, D. Zhuang, M. Li, J. Morris Chang, “Epi-Curriculum: Episodic Curriculum Learning for Low- Resource Domain Adaptation in Neural Machine Translation,” IEEE Transactions on Artificial Intelligence, Vol. 5, No. 12, pp. 6095–6108, Dec. 2024.
3. “Machine Translation Advancements of Low-Resource Indian Languages by Transfer Learning,” in Proceedings of the Ninth Conference on Machine Translation (WMT 2024), November 2024. — This includes work on low-resource Indian languages (Assamese, Manipuri, Khasi, Mizo) using transfer learning. Statistical Machine Translation
4. “Efficient Incremental Training using a novel NMT-SMT hybrid framework for translation of low- resource languages,” Kumar Bhuvaneswari, Murugesan Varalakshmi, Sep 2024.
5. IndiTranslate: Bridging Language Barriers in India; A. Rao, S. Patil, K. Reddy;2024
6. Hybrid Rule-Based and Neural Model for Regional Dialect Translation; M. Deshmukh, S. Patel;2023.