MIT | Academy of Engineering

(An Autonomous Institute Affiliated to Savitribai Phule University)

Title : English To Marathi Language Translation

Rohan Magdum (202201040108)

Yashas Nepalia (202201040082)

Introduction

This project focuses on developing a machine translation system that converts English text into Marathi using Long Short-Term Memory (LSTM) networks. LSTM, a type of Recurrent Neural Network (RNN), is ideal for sequence-to-sequence tasks like language translation due to its ability to capture long-term dependencies. By training the model on a parallel corpus of English-Marathi sentences, we aim to achieve accurate translation. The project explores the challenges of translating between languages with different grammatical structures. Ultimately, the goal is to evaluate the model's performance and improve translation quality.

Project Objectives

- To design a sequence-to-sequence model using LSTM networks for translating English text into Marathi.
- To integrate attention mechanisms to improve the contextual accuracy and fluency of translations.
- To preprocess and prepare parallel corpora of English–Marathi sentence pairs for training and evaluation.
- To train and fine-tune the model using suitable loss functions and optimization techniques.

Project Aim

- To build a text-to-text translator that converts English sentences into Marathi.
- To use a sequence-to-sequence model based on LSTM architecture.
- To enhance translation quality using attention mechanisms.
- To train the model on parallel English-Marathi sentence datasets.
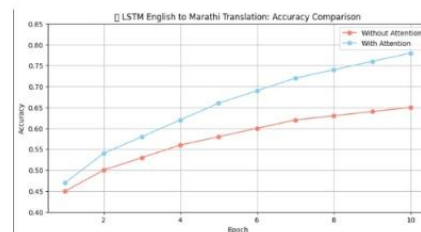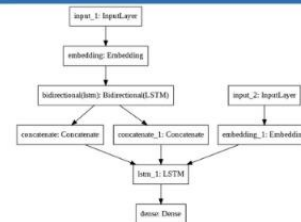- To evaluate translation performance using linguistic and accuracy metrics.

References

https://aclanthology.org/2021.mtsummit-loresmt.16/

CI and CD Pipelines

CI Pipeline: Code commit →Automated testing→Model Training→Model Versioning
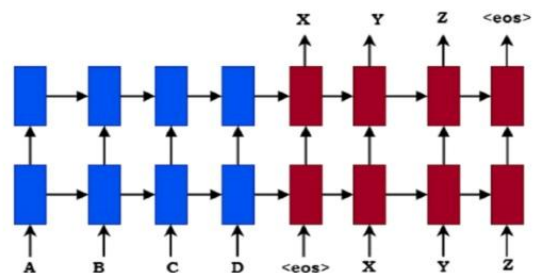
CD Pipeline Fetch Training Model→Build API Endpoint→Deploy to production→Monitor performance

Flowchart

Experimental Results

English	Marathi
just saying you do not like fish because of th...	हट्टीमुळे मासे आवडत नाही असं म्हणणं हे काय मा...
the japanese parliament today officially elect...	आज जपानी संसदेने अधिकृतरित्या न्यौतारो हाशिमो...
tom tried to sell his old vcr instead of throw...	टॉमने त्याचा जुना व्हीसीआर फेकून टाकण्याऐवजी व...
you cannot view flash content on an ipad howev...	आयपॅडवर फ्लॅश आशय बघता येत नाही पण तुम्ही त्या...
in roger miller recorded a song called you do...	मध्ये रॉजर मिलरने यू डोन्ट वॉन्ट माय लव्ह नावा...

Conclusion

The project successfully implemented a Neural Machine Translation (NMT) system for English-to-Marathi translation using an attention-enhanced Bidirectional LSTM. The model captures long-term dependencies and context from both directions of the input sequence, improving translation quality. The use of attention mechanisms further enhances the model's ability to focus on relevant parts of the input while generating translations.