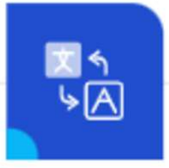
Neural Machine Translation System

A comprehensive exploration of neural machine translation techniques for English to Marathi using deep learning.

Team Members:

- (202201040124) Shriram Savant
- (202201040126) Suraj Didwagh
- (202201040138) punit Kawadkar
- (202201040183) Prathamesh Galugade

Introduction to NMT Systems



Objective of NMT System

Develop a Neural Machine Translation (NMT) system to translate English sentences to Marathi using a Bidirectional Long Short-Term Memory (LSTM) model with an attention mechanism.

Importance of NMT

Enables accurate translation for low-resource languages like Marathi, making information accessible to broader audiences.

Role of Deep Learning

Leverages deep learning techniques to capture contextual nuances in translation, enhancing understanding of language intricacies.

Significance of Attention Mechanism

Improves translation quality for longer sentences by focusing on relevant words, thus maintaining contextual integrity.

Overview of English-Marathi Dataset

- **Dataset Size: 41,028 pairs**

  The dataset consists of a total of **41,028 sentence pairs**, providing a robust foundation for training and testing models.

- **Preprocessing Steps**

  Key preprocessing steps include adding **'sos'** and **'eos'** tokens, crucial for sequence prediction tasks.

- **Tokenization and Padding**

  Sentences were **tokenized** and **padded** to a fixed length of **36 tokens**, ensuring uniform input size for the model.

- **Train-Test Split**

  The data was split into **90% training** and **10% testing** sets, allowing for effective validation of the model's performance.

- **Example Sentence Pair**

  An example from the dataset: *English:* 'I sleep with the lights on' *Marathi:* '□□□□□□□□□□□□□□□□□□□□'

Neural Translation Architecture

01 **Encoder: Bidirectional LSTM**

Utilises 512 units to effectively capture context from both directions, improving understanding of input sequences.

02 **Attention Layer:**

Focuses on relevant encoder outputs by computing weights, enhancing the translation quality through selective attention.

03 **Decoder: LSTM with Attention**

Employs a 512-unit LSTM paired with attention to generate accurate Marathi translations from encoded information.

04 **Embedding: 128-dimensional**

Utilises 128-dimensional word embeddings for both English and Marathi, facilitating better semantic representation of words.

05 **Contextual Understanding**

The bidirectional nature of the encoder aids in understanding the full context of sentences, crucial for accurate translation.

06 **Translation Quality**

The combination of LSTM and attention mechanisms significantly enhances the quality of translations between languages.

Training Details for NMT

Optimizer: Adam

The **Adam** optimizer is used for updating the weights during training, known for its efficiency and effectiveness in handling sparse gradients.

Loss Function: Categorical Crossentropy

Categorical Crossentropy is employed as the loss function, suitable for multi-class classification tasks like language translation.

Training Epochs: 10

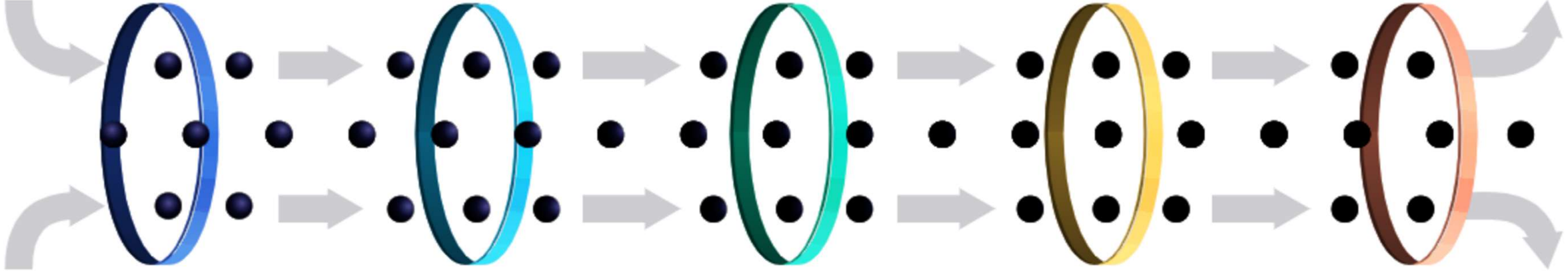
The model is trained for **10 epochs**, allowing it to learn from the data multiple times for better accuracy.

Batch Size: 64

A **batch size of 64** is used, which determines the number of training samples processed before the model's internal parameters are updated.

Inference Method: Greedy Decoding

During inference, separate encoder and decoder models are employed, utilising **greedy decoding** for generating translations efficiently.



Performance Results Overview

01

Performance analysis conducted

Tested on 4,103 sentences, yielding promising results in translations.

02

Sample Translations provided

Examples of translations from English to Marathi demonstrate system capability.

03

Translation Accuracy emphasis

Focus on the accuracy and reliability of the predicted translations.

04

Diversity of Input Sentences

The test included a variety of sentence structures and contexts for robust evaluation.

05

Encoder-Decoder Framework

Utilises an encoder-decoder architecture for effective translation processes.

06

Attention Mechanism significance

Incorporates attention mechanisms to improve translation quality and coherence.

Challenges in Neural Translation

Rare words pose challenges

Translating uncommon vocabulary requires nuanced understanding.

Idiomatic expressions complicate

Idioms may not have direct translations, leading to confusion.

Longer sentences need focus

Extended sentences can overwhelm attention and processing.

Contextual understanding is key

Effective translation relies on grasping the context fully.

Language structure differences

English and Marathi have distinct grammatical structures affecting translation.

Neural Machine Translation Insights

Neural Machine Translation is crucial for language conversion.

It employs complex algorithms to convert languages accurately and efficiently.

Encoder-Decoder architecture facilitates translation.

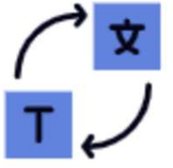
This structure encodes the source language and decodes it into the target language.

Attention mechanisms enhance translation results.

They allow the model to focus on relevant parts of the input sentence during translation.

Visit the project at GitHub for more details.

Explore the complete codebase and documentation for implementation.



English to Marathi translation is an application of this technology.

This specific translation task benefits from the outlined methodologies.

Enhancing English to Marathi Translation

NMT System Enhancements

The NMT system significantly improves translation accuracy for English to Marathi, providing more reliable results.

Attention Mechanisms Role

Attention mechanisms enhance the understanding of longer sentences, allowing for better contextual translation.

Future Advancements Potential

Future advancements may further improve performance and expand capabilities in translation technologies.

CONTACT US

Contact Information for Collaboration

We welcome discussions and partnerships related to English to Marathi translation using advanced deep learning techniques.