PROJECT GUIDE: MS DIPTEE GHUSSE

GROUP MEMBERS:
SHRUTI DEOKAR
RIYA GARASANGI

MIT | Academy of Engineering

Comparative Study of Encoder-Decoder Architectures with Attention Mechanisms for Grammaar Correction

Introduction

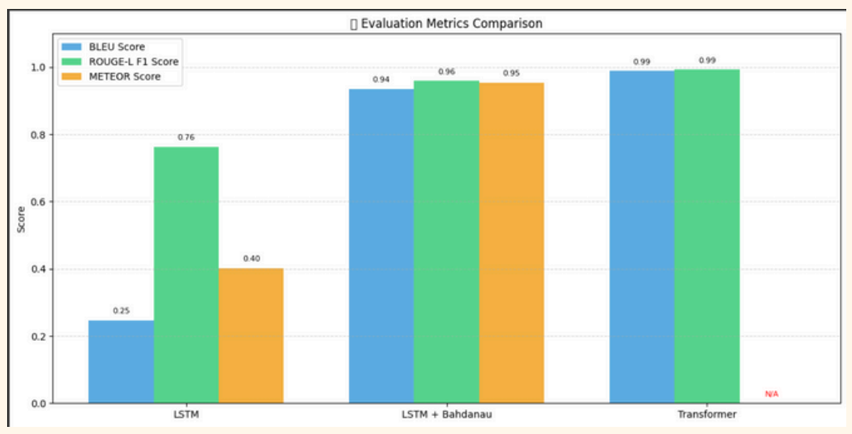
This project compares three encoder-decoder models for grammar correction: LSTM without attention, LSTM with Bahdanau attention, and Transformer with self-attention. The goal is to evaluate how attention mechanisms affect accuracy and efficiency.

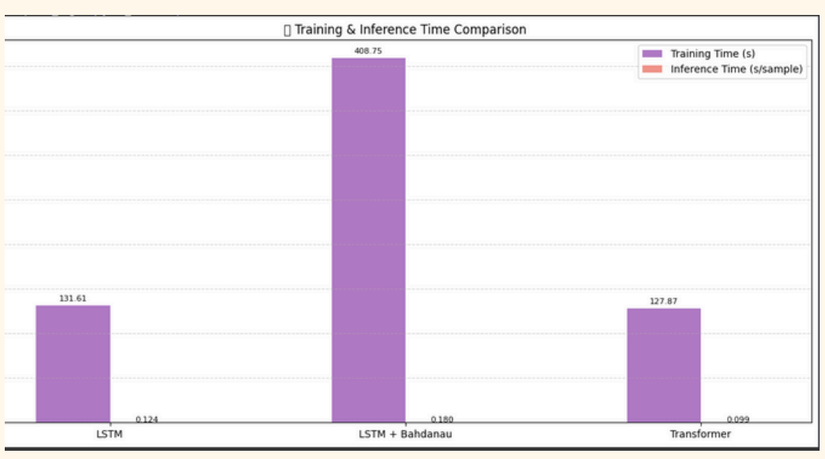
Dataset

- Dataset Name: Grammar Correction
- Link: https://www.kaggle.com/datasets/satishgunjal/grammar-correction
- Dataset Size: 1,560 rows × 4 columns.
- Preprocessing : Tokenization

Abstract

We compare three models for grammar correction: LSTM, LSTM with Bahdanau attention, and Transformer with self-attention. Trained on sentence pairs and evaluated using BLEU and ROUGE, results show that attention-based models—especially Transformers—achieve better accuracy and grammar correction performance.





Model	BLEU	ROUGE-L	METEOR	Training Time (s)	Inference Time (s/sample)
LSTM	0.247	0.7627	0.4017	131.61	0.124
LSTM + Bahdanau	0.9352	0.9596	0.9537	408.75	0.1798
Transformer	0.9886	0.9933	N/A	127.87	0.0993

Methodology

• Dataset: Custom dataset of incorrect and corrected sentence pairs, preprocessed with tokenization.
• Models:
1. LSTM Without Attention
2. Bahdanau Attention: Seq2seq with dynamic attention.
3. Transformer: Self-attention model for long dependency handling.

Analysis

- LSTM alone struggles with long-term dependencies → lower BLEU & ROUGE.
- LSTM + Attention improves performance significantly (as shown by BLEU jump from 0.24 to 0.93).
- Transformer architecture (self-attention + positional encoding) dominates in all areas, offering both accuracy and speed.