Comparative Study of Encoder-Decoder Architectures for English Grammar Correction

MIT Academy of Engineering
(An Autonomous Institute)

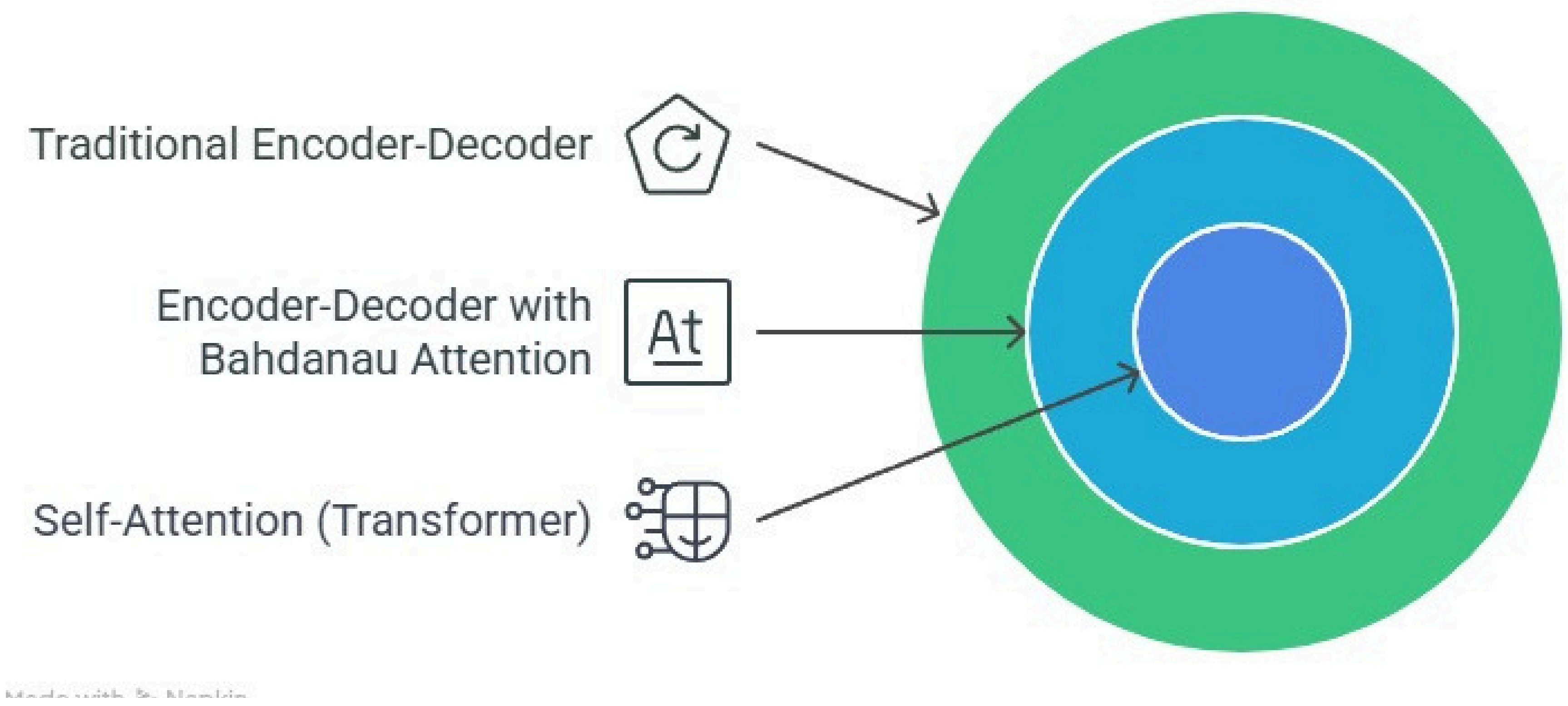
Md. Yaseen Alam
Vaibhav Jadhav
Om Borle
Vidya Bingi
(T.Y. B.Tech Comp Engg)

Abstract

This project explores English grammar correction using deep learning models: LSTM/GRU without attention, Bahdanau attention, and Transformer with self-attention. We implemented and compared these models using a custom dataset of grammatically incorrect and corrected sentence pairs. The models were evaluated using BLEU score and accuracy. Results show that the Transformer model outperforms others in both speed and quality, making it ideal for real-time grammar correction applications.

Dataset

- Collected custom dataset of grammatically incorrect and corrected sentence pairs.
- Preprocessing: tokenization, lowercasing, padding.
- Sample:
- Incorrect: "The most most powerful superhero is Superman."
- Corrected: "The most powerful superhero is Superman."



Introduction

This project investigates the effectiveness of three encoder-decoder architectures in correcting ungrammatical English sentences: LSTM/GRU (without attention), Bahdanau attention, and Transformer (with self-attention). We aim to understand how attention mechanisms impact grammar correction and identify the best-performing model.

METHODOLOGY

- Dataset: Custom dataset of incorrect and corrected sentence pairs, preprocessed with tokenization, lowercasing, and padding.
- Models:

1. LSTM/GRU: Basic sequence-to-sequence model.
2. Bahdanau Attention: Seq2seq with dynamic attention.
3. Transformer: Self-attention model for long dependency handling.

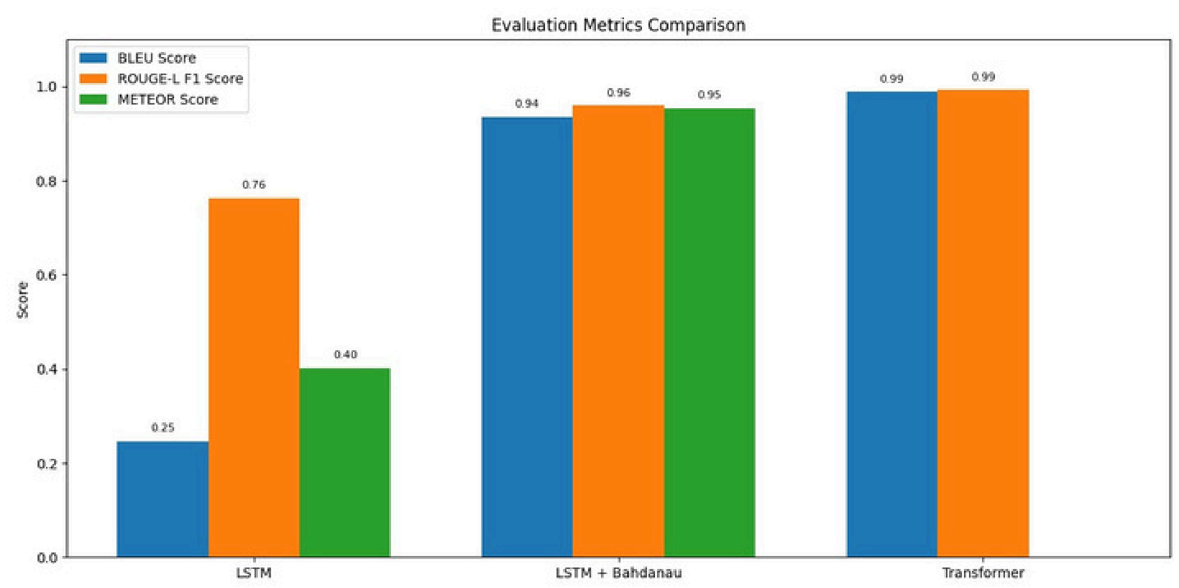
- Training:
  - Optimizer: Adam
  - Loss: Categorical Crossentropy
  - Metrics: BLEU, ROUGE, METEOR and Accuracy.

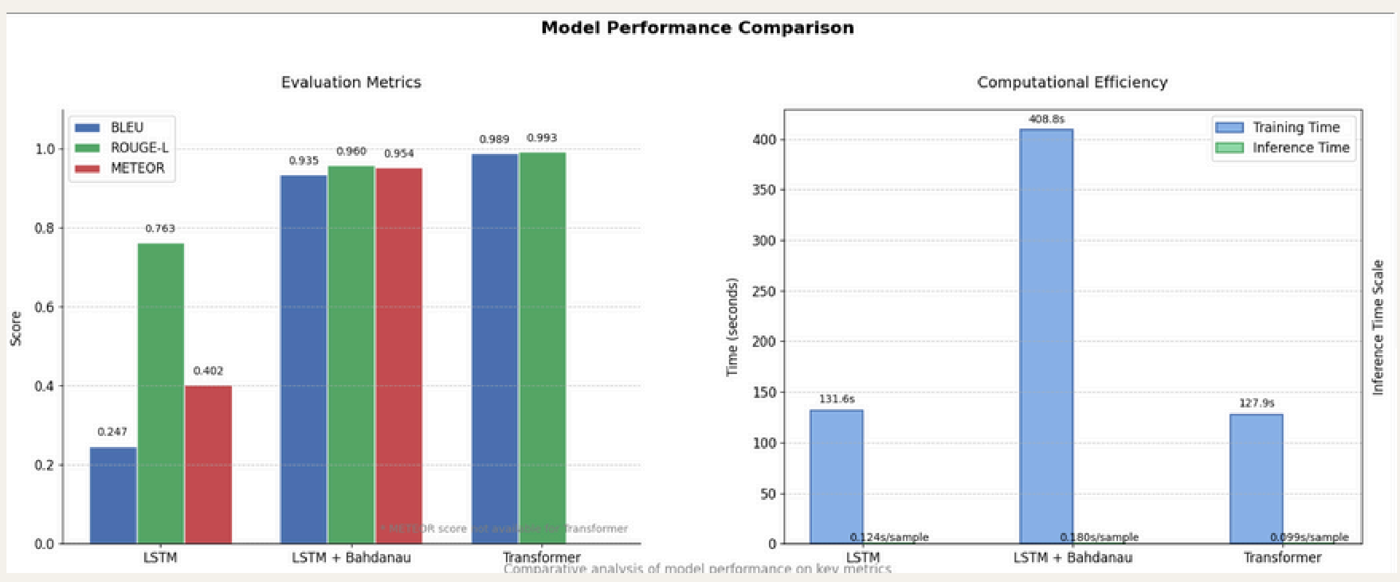
RESULTS

Criteria	LSTM/GRU (No Attention)	Attention (Bahdanau/Luong)	Transformer (Self-Attention)
Accuracy / BLEU	0.2470	0.9352	0.9886
ROUGE	0.7627	0.9596	0.9933
METEOR	0.4017	0.9537	NA
Training Time	130.61 seconds	408.75 seconds	127.87 seconds
Inference Speed	0.1240 sec/sample	0.1798 sec/sample	0.0993 sec/sample

Analysis

- Performance: Transformer achieves the highest accuracy (BLEU: 0.99), followed by LSTM + Attention, showing attention significantly improves results. LSTM without attention performs the worst.
- Efficiency: Transformer is the fastest to train and infer, while LSTM + Attention is the slowest.
- Insight: Self-attention mechanisms greatly enhance grammar correction accuracy and speed.





Conclusion

Transformer models with self-attention significantly outperform LSTM/GRU and Bahdanau attention models, offering faster and more accurate grammar correction. Future work will focus on expanding the dataset, fine-tuning pretrained models like T5 or BERT, and exploring deployment options.

Acknowledgement

Thanks to our professors and peers at MIT Academy of Engineering for their support throughout this project.