MIT | Academy of Engineering
(An Autonomous Institute Affiliated to Savitribai Phule University)

Group Members:
Sapna Dahikamble_202201070065
Supriya Maskar _202201040049
Manjiri Netankar _202201040206

PARAPHRASE GENERATION: COMPARING ATTENTION MECHANISMS

1 INTRODUCTION

This project aims to explore and compare various attention mechanisms used in paraphrase generation tasks. Paraphrasing is the process of generating alternative expressions that convey the same meaning, and it plays a crucial role in applications such as:

- Question answering
- Text summarization
- Chatbots and dialogue systems

2 OBJECTIVE

To evaluate and compare the performance of three attention mechanisms:

- Bahdanau Attention
- Luong Attention
- Transformer Self-Attention

within a sequence-to-sequence (Seq2Seq) paraphrase generation framework.

3 DATASET

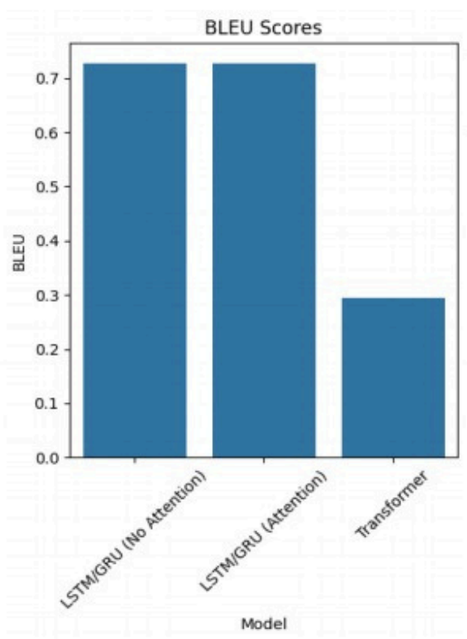
Reduced Quora Question Pairs (20,000 samples)

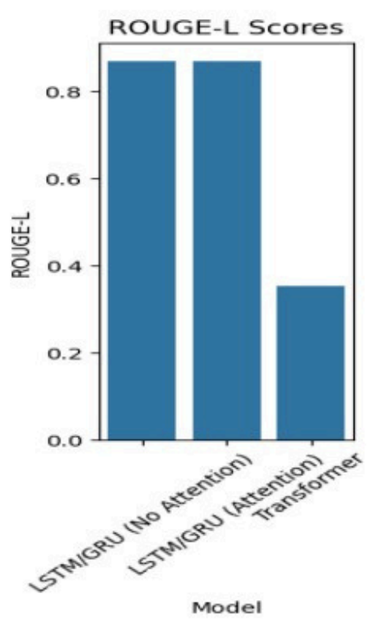
- Contains sentence pairs: input and target (paraphrases)
- Preprocessed for training Seq2Seq models
- Aligned for encoder-decoder structure with attention
- Used to evaluate semantic similarity in paraphrase generation

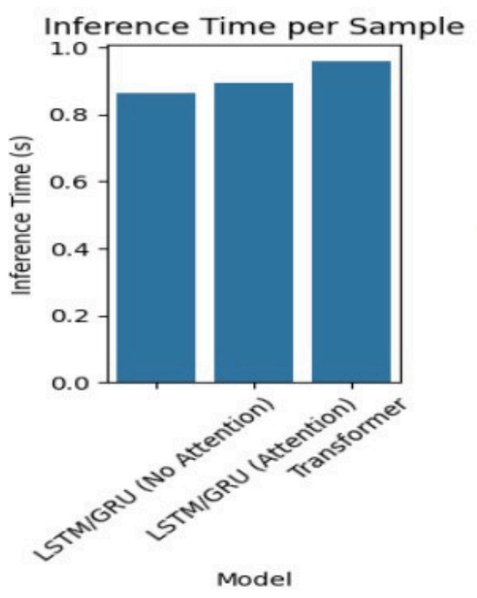
4 MODEL ARCHITECTURES

All models follow the encoder-decoder structure, with different attention modules:

- Bahdanau Attention: Additive attention mechanism
- Luong Attention: Multiplicative attention with global/local scope
- Transformer: Uses self-attention without recurrence, allowing parallelization







5 EVALUATION METRICS

- BLEU Score: Measures precision of n-gram matches
- ROUGE Score: Measures recall of overlapping sequences
- Accuracy: % of predicted paraphrases matching the ground truth

6 WORKING PRINCIPLE

- Encoder converts input sentence into hidden states
- Attention Mechanism determines relevance of input words during decoding
- Decoder generates output word-by-word, guided by attention weights

7 CONCLUSION

- Transformer-based models outperform traditional attention mechanisms
- Bahdanau attention performs better than Luong in this setup
- Attention plays a vital role in capturing context for better paraphrasing