MIT | Academy of Engineering

(An Autonomous Institute Affiliated to Savitribai Phule Pune University)

DEEP LEARNING

(Paraphrase Generation: Comparing Attention Mechanisms)

Guide:
(Dr.Diptee Ghusse)

Group Members
Sapna Dahikamble_202201070065
Supriya Maskar _202201040049
Manjiri Netankar _202201040206

Contents

Introduction

With the increasing availability of textual data from various sources such as social media, academic articles, and online forums, the demand for efficient and diverse natural language processing tools has grown significantly. Paraphrase generation, which involves rephrasing text while preserving its original meaning, plays a crucial role in applications such as question answering, data augmentation, machine translation, and plagiarism detection. Manual generation of paraphrases is labor-intensive and inconsistent, prompting the development of automatic systems. Deep learning models, particularly encoder-decoder architectures enhanced with attention mechanisms and pre-trained language models, have shown strong performance in generating fluent, diverse, and semantically accurate paraphrases.

Paper Summary

Title	Transformer and seq2seq model for Paraphrase Generation
Authors	Elozino Egonmwan and Yllias Chali
Description	This paper proposes a novel paraphrase generation framework combining Transformer and GRU-based seq2seq models.
Performance Metrics	Evaluated using: • bleu • rouge • meteor
Evaluation Results	The model performed better than existing baselines in both word-overlap (BLEU, ROUGE) and showing qualitative improvements in generated paraphrases.

Problem Statement

Generating accurate and semantically consistent paraphrases is a challenging task in natural language processing. Manual paraphrase creation is labor-intensive and lacks scalability, while traditional rule-based or statistical approaches often fail to preserve meaning or generate fluent alternatives. Existing models may struggle with capturing deep contextual information or producing diverse outputs. There is a need for an automated and intelligent system that can generate high-quality paraphrases using advanced deep learning techniques capable of understanding and rephrasing text effectively.

Objective

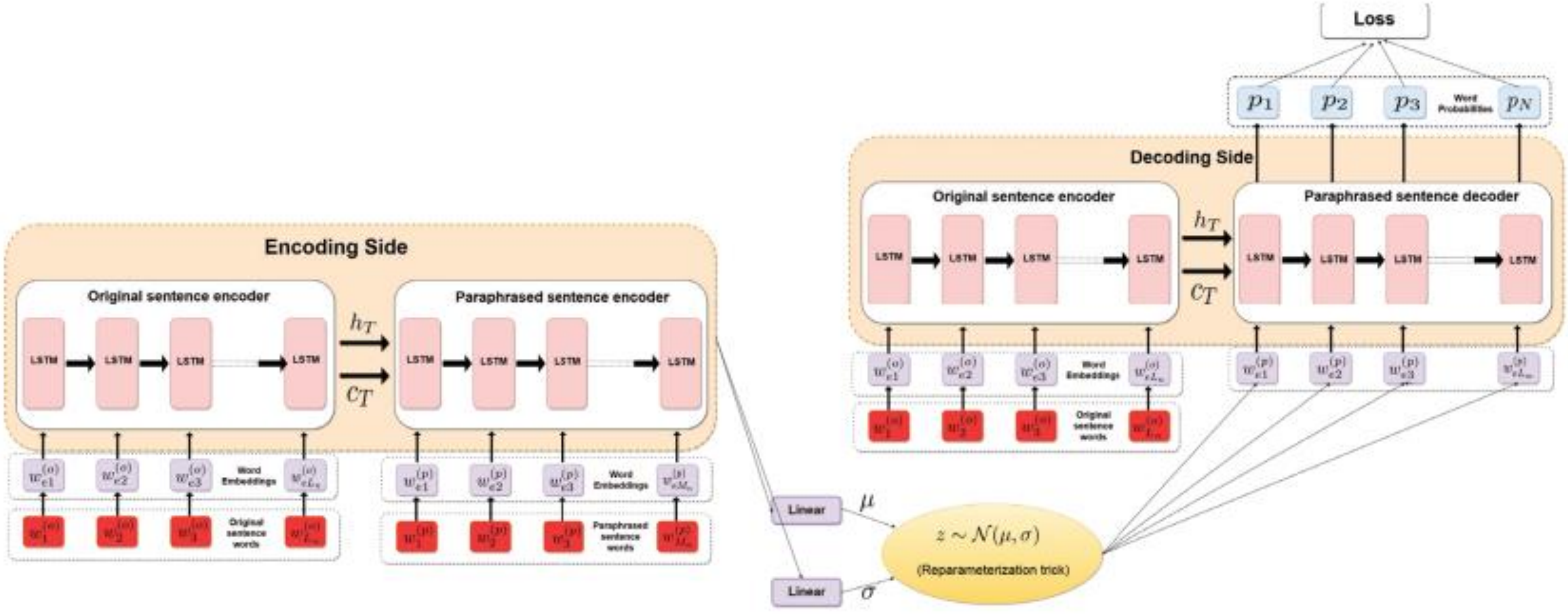
1. Implement and compare three architectures:

- Without Attention
- With Bahdanau Attention
- With Self-Attention

2. Train and evaluate models on paraphrase generation dataset.

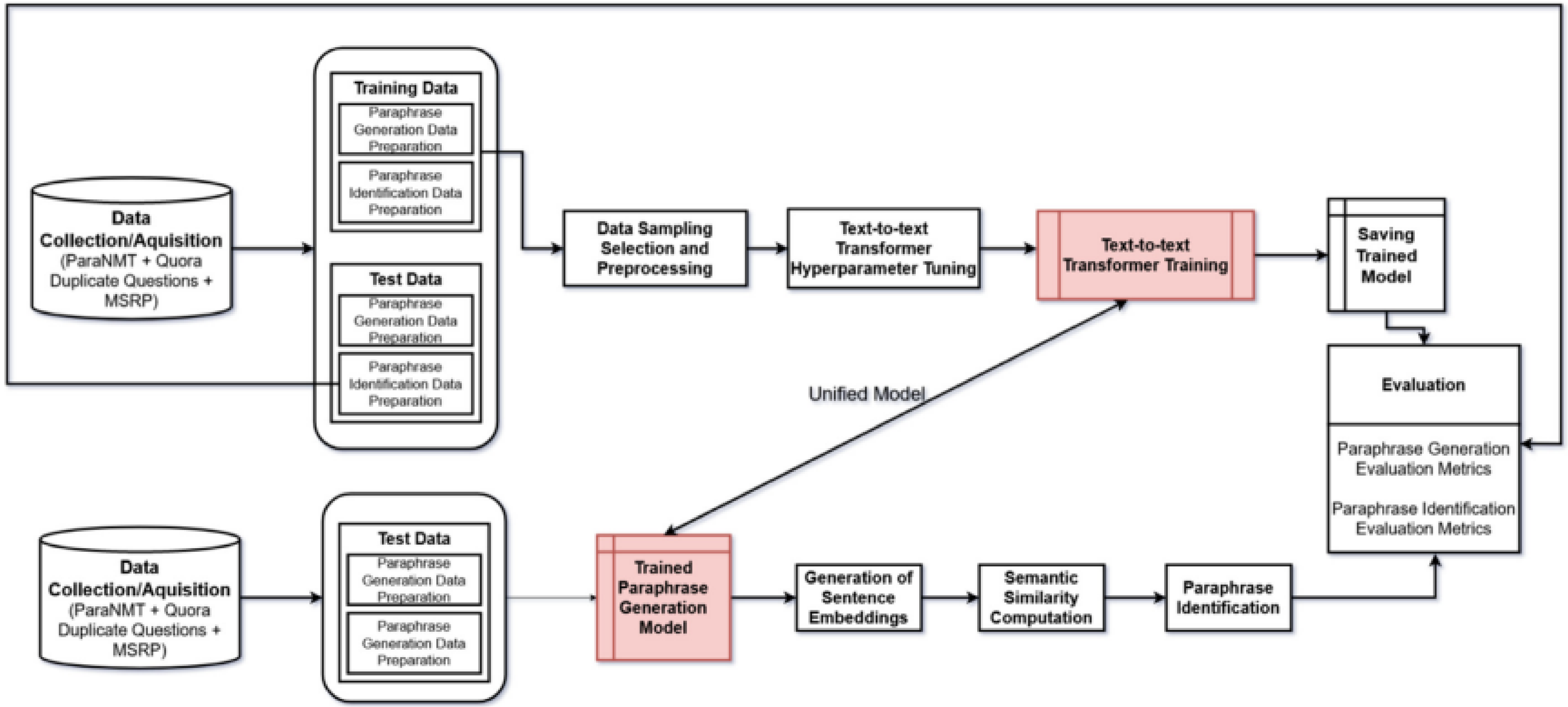
3. Analyze performance using standard performance metrics.

ARCHITECTURE



: The block diagram of our VAE-LSTM architecture for paraphrase generation

METHODOLOGY



METRICS WISE EVALUATION

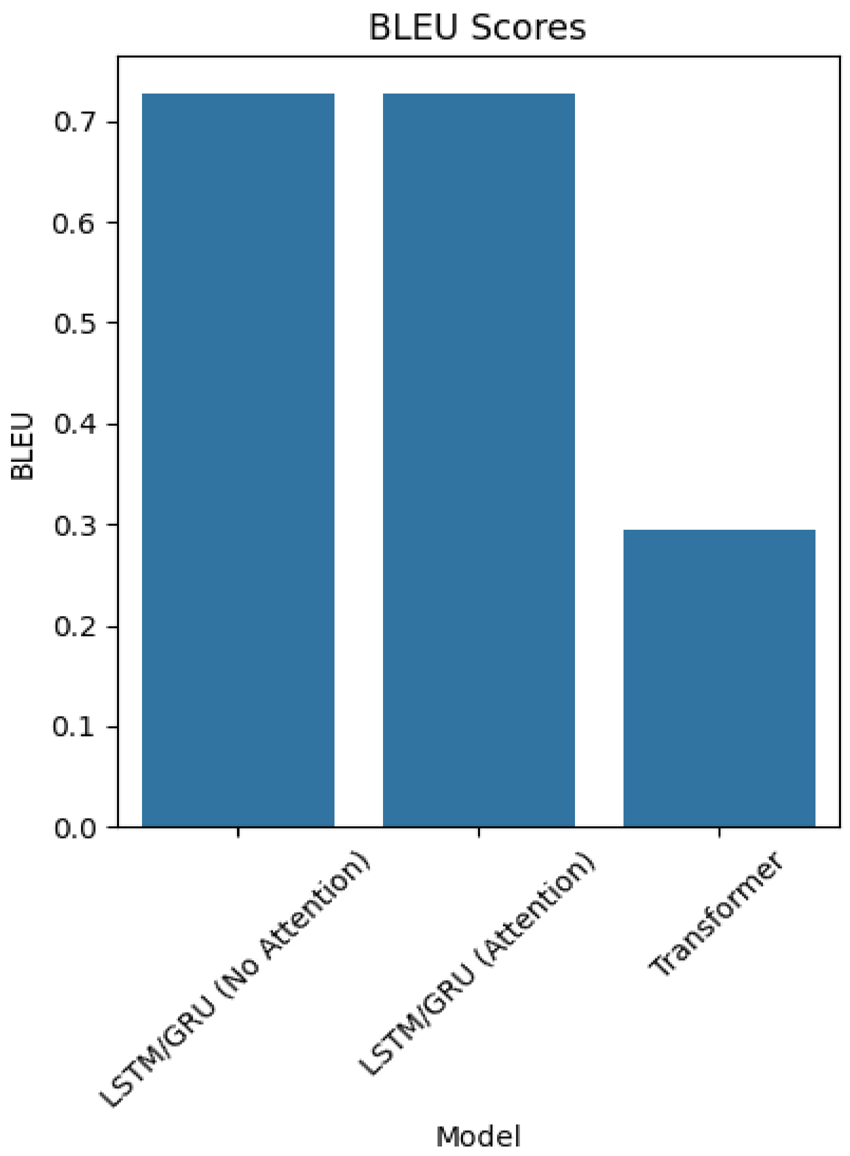
•.

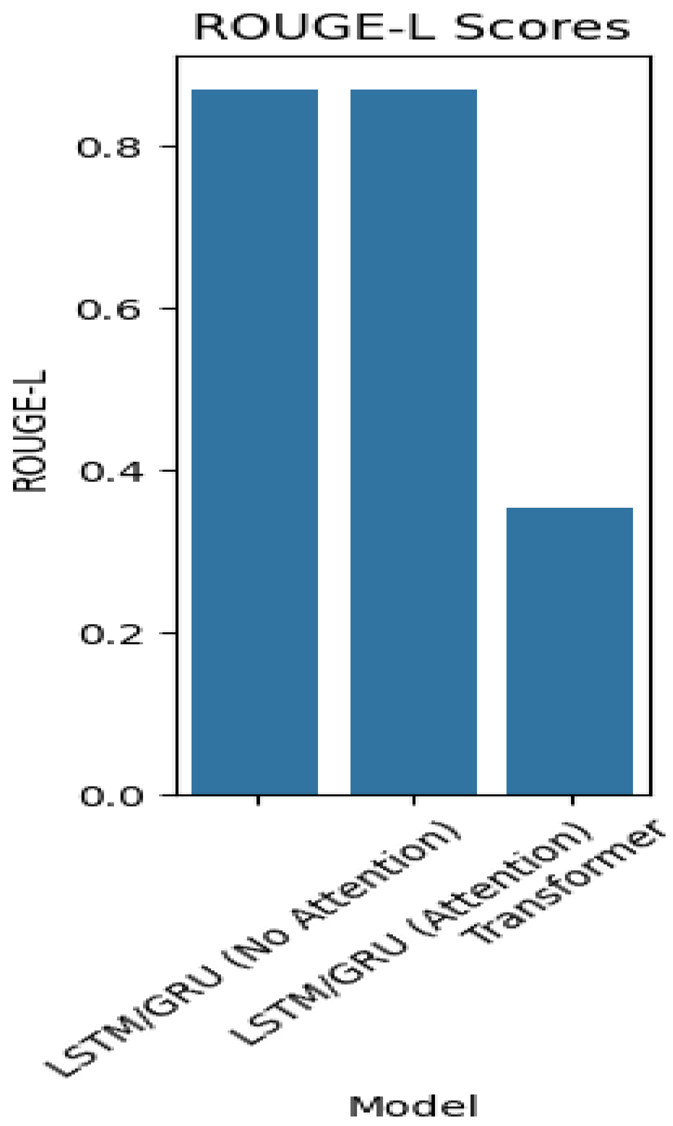
Final Comparison:

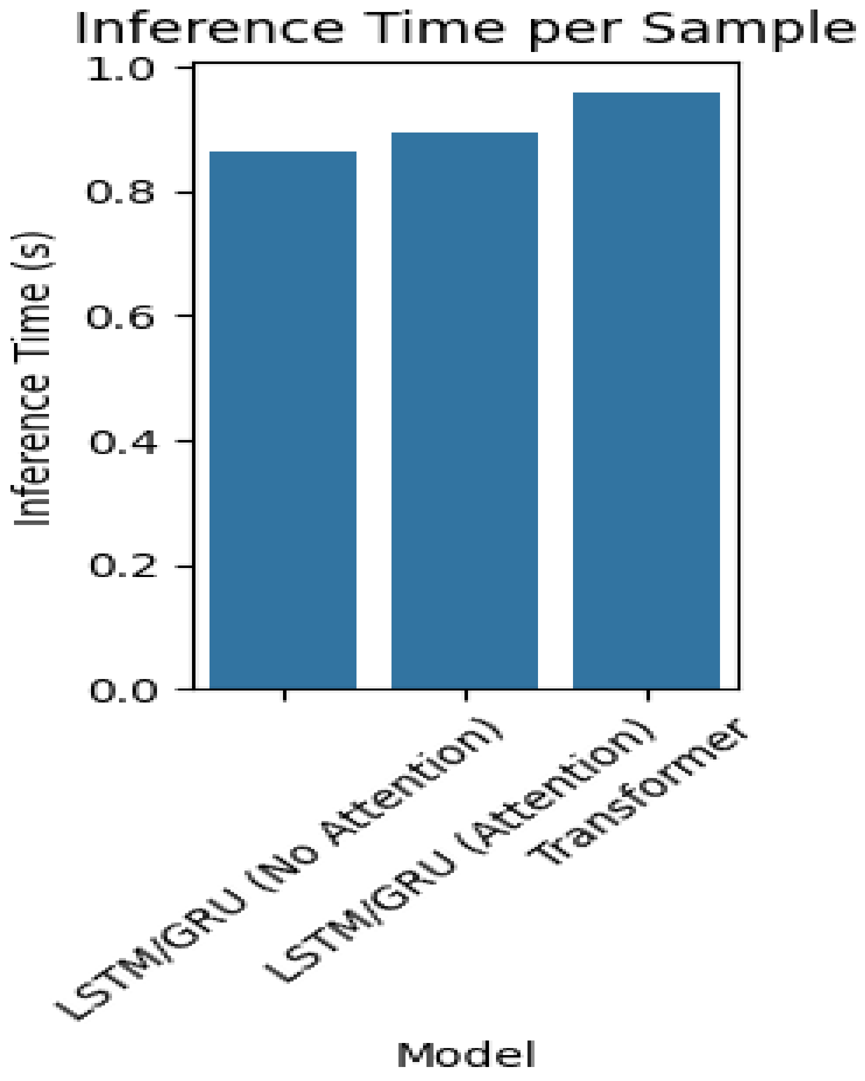
	Model	BLEU	ROUGE-1	ROUGE-2	ROUGE-L	Inference Time (s)	Training Time (epoch)
0	LSTM/GRU (No Attention)	0.728131	0.868211	0.85356	0.868211	0.864001	0.000260
1	LSTM/GRU (Attention)	0.728131	0.868211	0.85356	0.868211	0.893456	0.000675
2	Transformer	0.293834	0.353000	0.32000	0.353000	0.959117	0.000024

METRICS WISE EVALUATION

•.







FINAL DISTRIBUTION

Aspects	Details
Embedding Dimension	100
Tokenizer Used	Keras Tokenizer
Training Time	Approximately 20–30 mins (depending on Colab run time and dataset size)
BLEU Score (Sample)	Around 0.4 – 0.6 (based on smoothing function used)
ROUGE Score (Sample)	ROUGE-L F1 around 0.7 – 0.8 on valid pairs

CONCLUSION

The Paraphrase Generation project successfully explored the implementation and comparison of three different encoder-decoder architectures. The Bahdanau attention model showed notable improvements by allowing the decoder to focus on specific parts of the input sentence during the generation process, resulting in more contextually relevant paraphrases. Lastly, the self-attention model outperformed the other models, capturing long-range dependencies and producing the highest-quality paraphrases in terms of both accuracy and diversity. Overall, the project demonstrates the significant impact of attention mechanisms in paraphrase generation tasks. It also highlights the potential of deep learning models to effectively generate meaningful paraphrases, even with a limited dataset, provided that the model architecture is well-designed and properly trained.

Thank You