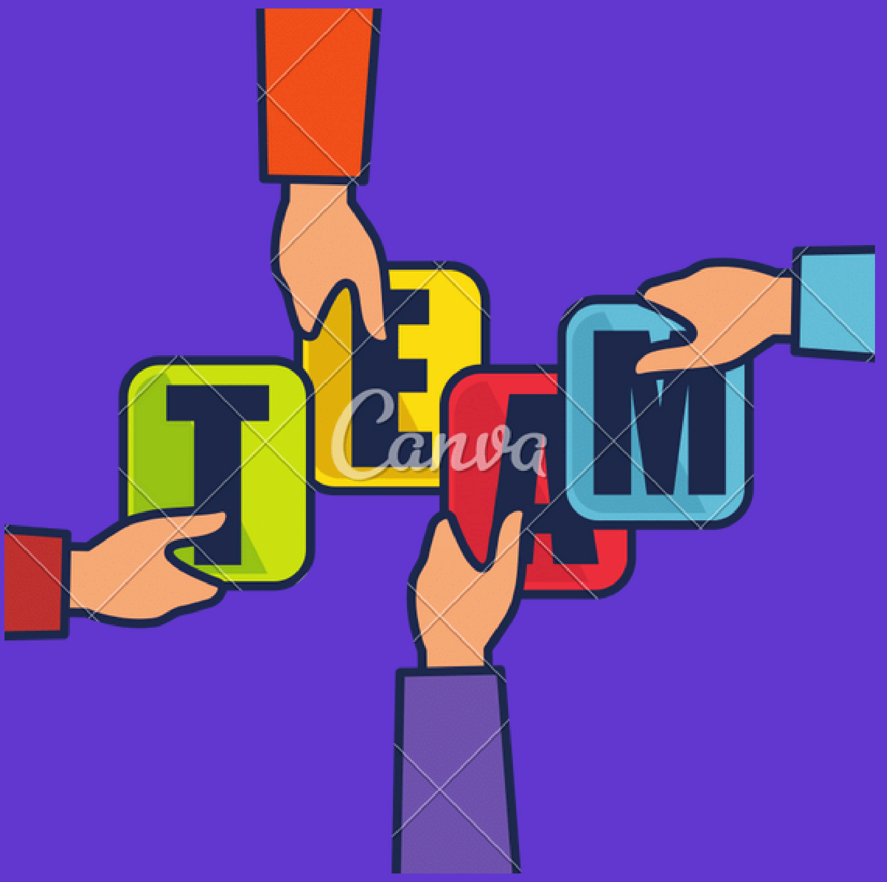
MIT | Academy of Engineering
(An Autonomous Institute Affiliated to Savitribai Phule Pune University)

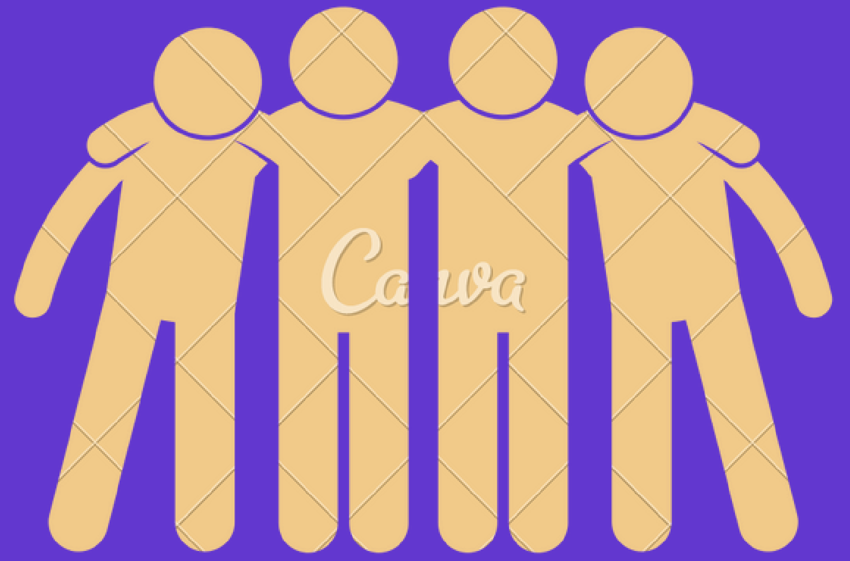
Deep Learning Lab Mini Project

Grammar Correction

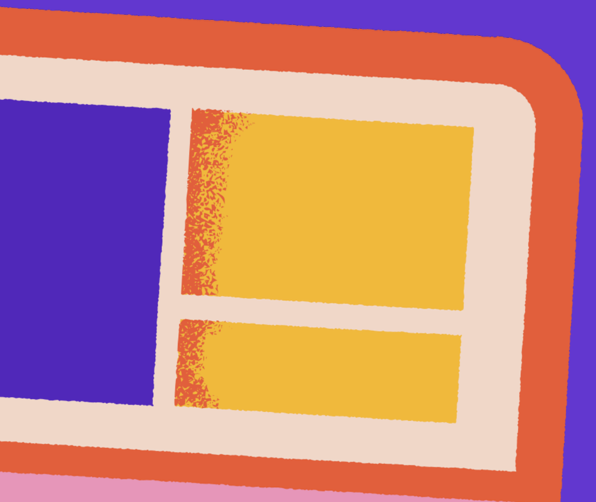
Course Instructor : Diptee Ghusse



Our Team



Mayur Kapgate	PRN : 202201040065
Anirudha Gapat	PRN : 202201040067
Ganesh Atre	PRN : 202201040068



Content

- Introduction
- Paper Summary
- Dataset description
- Model diagrams & architecture
- Metric-wise performance comparison
- Graphs
- Final analysis table
- Conclusion

Introduction

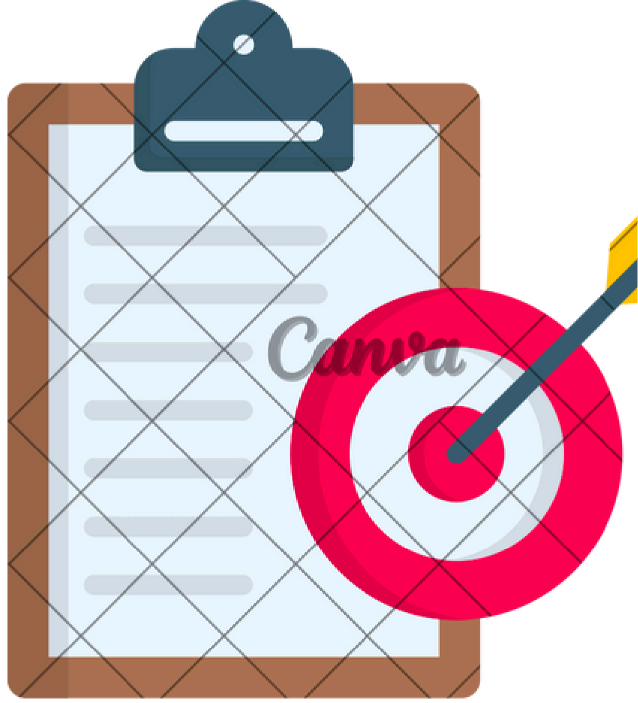
Grammatical Error Correction (GEC) can be considered a specialized form of paraphrase generation, where the objective is to produce a grammatically correct version of the original sentence by applying minimal and targeted edits.

Aim : To implement and compare different encoder-decoder architectures for the for the Grammaer Correction

We explore models:

- Without Attention (LSTM-based encoder-decoder)
- With Attention (Luong Attention mechanism)
- With Self-Attention (Transformer)

Paper Summary



Paper Title: GECToR – Grammatical Error Correction: Tag, Not Rewrite

Aim: To develop a fast, efficient, and high-performance grammatical error correction (GEC) system using a sequence tagging approach instead of traditional sequence-to-sequence (seq2seq) methods.

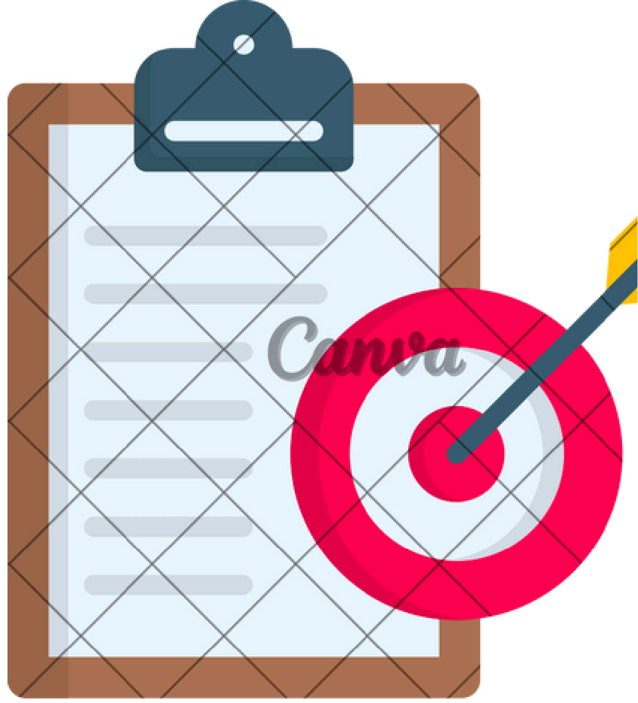
Objectives: Simplify GEC by transforming it into a sequence tagging task.

1. Improve inference speed and model interpretability.
2. Achieve state-of-the-art performance on standard GEC benchmarks.

Problem Statements:

1. Seq2seq GEC systems suffer from slow inference speeds and high computational demands.
2. Large amounts of training data are required for effective performance.
3. Lack of interpretability in corrections made by seq2seq models.

Paper Summary



- **Methodology:**

1.Model Architecture: Utilized a Transformer encoder) with token-level transformations for edits.

2.Token-Level Transformations: Designed custom transformations (e.g., KEEP, DELETE, APPEND, REPLACE) and grammatical-specific edits (e.g., CASE, MERGE, SPLIT).

3.Iterative Tagging: Applied corrections iteratively to handle dependencies between edits.

4.Evaluation: Tested on CoNLL-2014 and BEA-2019 datasets using precision (P), recall (R), and F0.5F0.5 scores.

Key Results:

•Achieved F0.5F0.5 scores of 65.3 (single model) and 66.5 (ensemble) on CoNLL-2014.

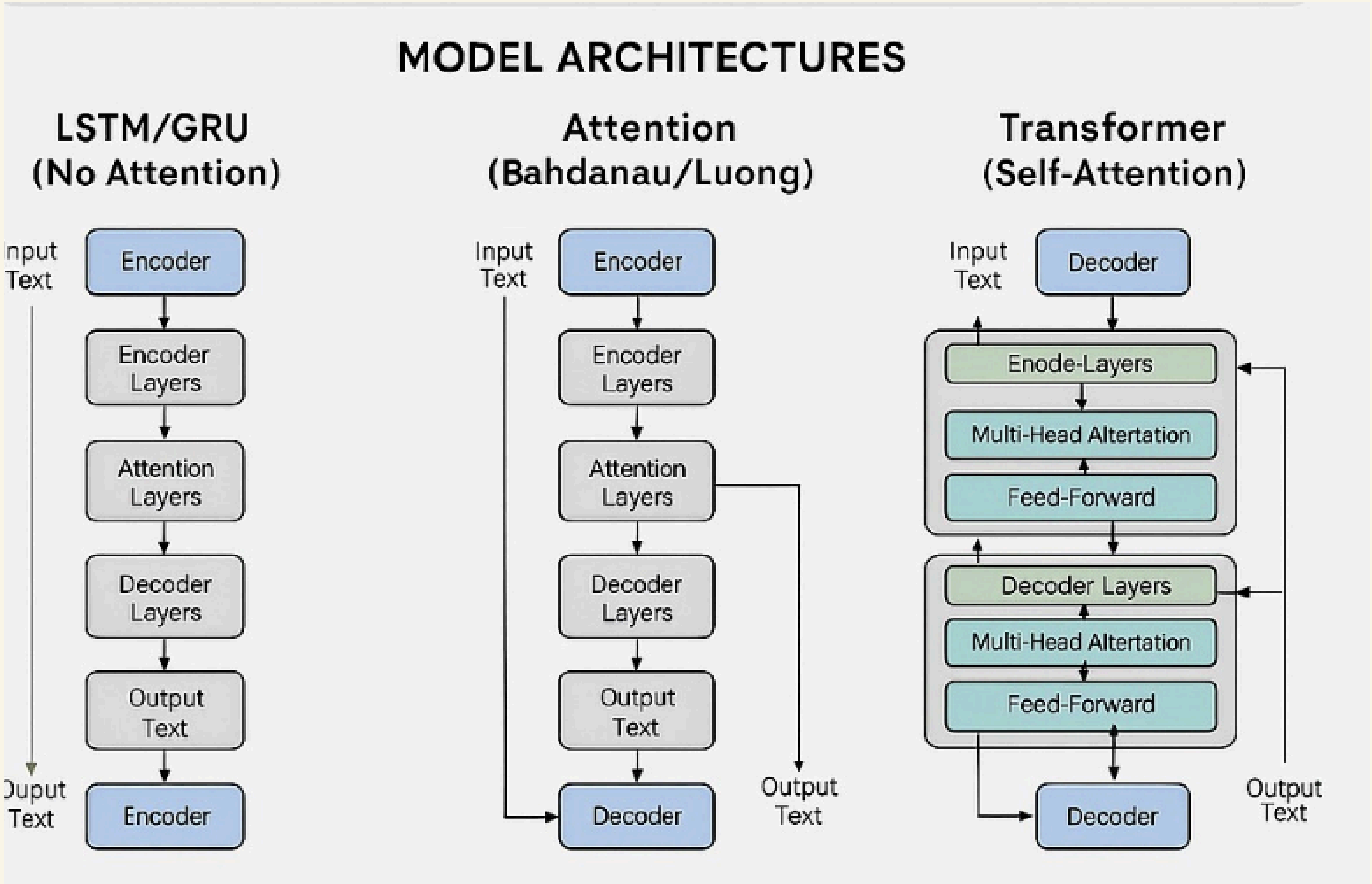
•Attained F0.5F0.5 scores of 72.4 (single model) and 73.6 (ensemble) on BEA-2019.

•Demonstrated up to 10x faster inference compared to seq2seq Transformer models.

Conclusion:

GECToR provides a simpler, faster, and more interpretable alternative to seq2seq GEC systems while achieving competitive performance. The use of token-level transformations and multi-stage training enables efficient and accurate grammatical error correction.

Diagram :


MODEL ARCHITECTURES
LSTM/GRU
(No Attention)
Input
Text
Encoder
Encoder
Layers
Attention
Layers
Decoder
Layers
Output
Text
Ouput
Text
Encoder
Attention
(Bahdanau/Luong)
Input
Text
Encoder
Encoder
Layers
Attention
Layers
Decoder
Layers
Output
Text
Decoder
Output
Text
Transformer
(Self-Attention)
Input
Text
Decoder
Enode-Layers
Multi-Head Altertation
Feed-Forward
Decoder Layers
Multi-Head Altertation
Feed-Forward
Decoder
Output
Text

Dataset Description

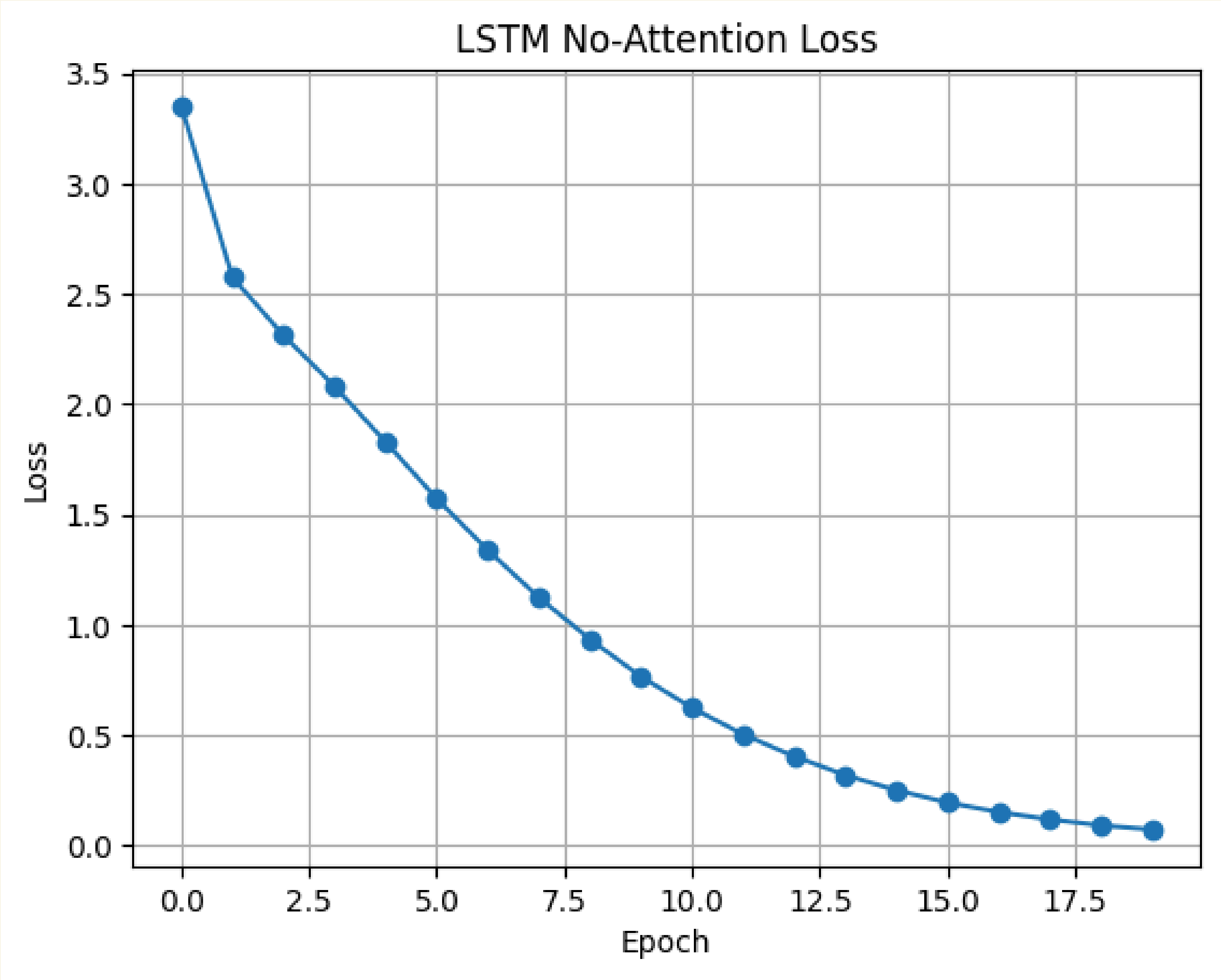
Dataset Name: Grammer Correction Dataset
Source: Kaggle

Dataset Size: Total Question Pairs: ~10000

Preprocessing Performed:
Text cleaning: Lowercasing
Tokenization and Padding for model input.

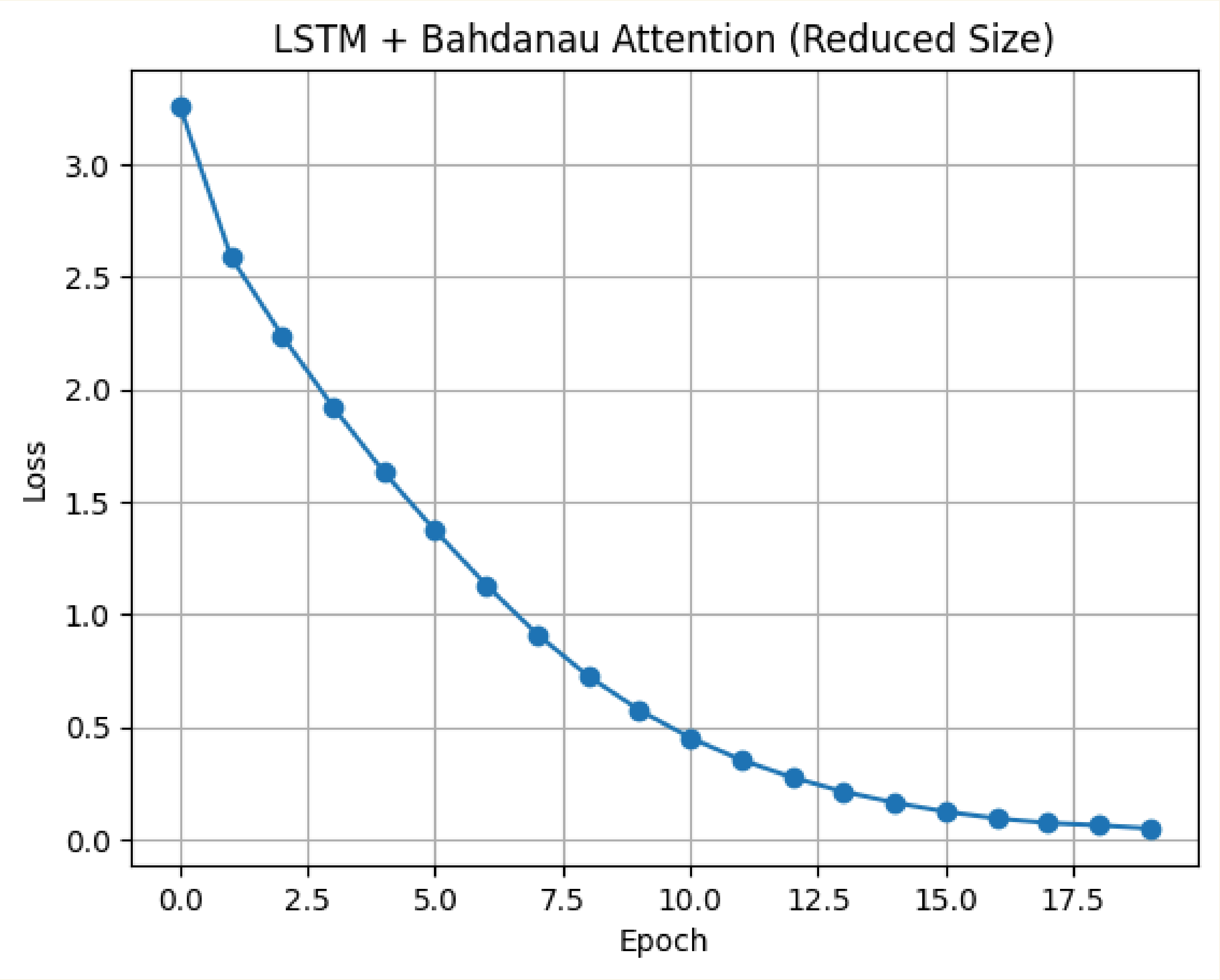
Graphs : Training Curves

- **Encoder-Decoder (LSTM) Without Attention**



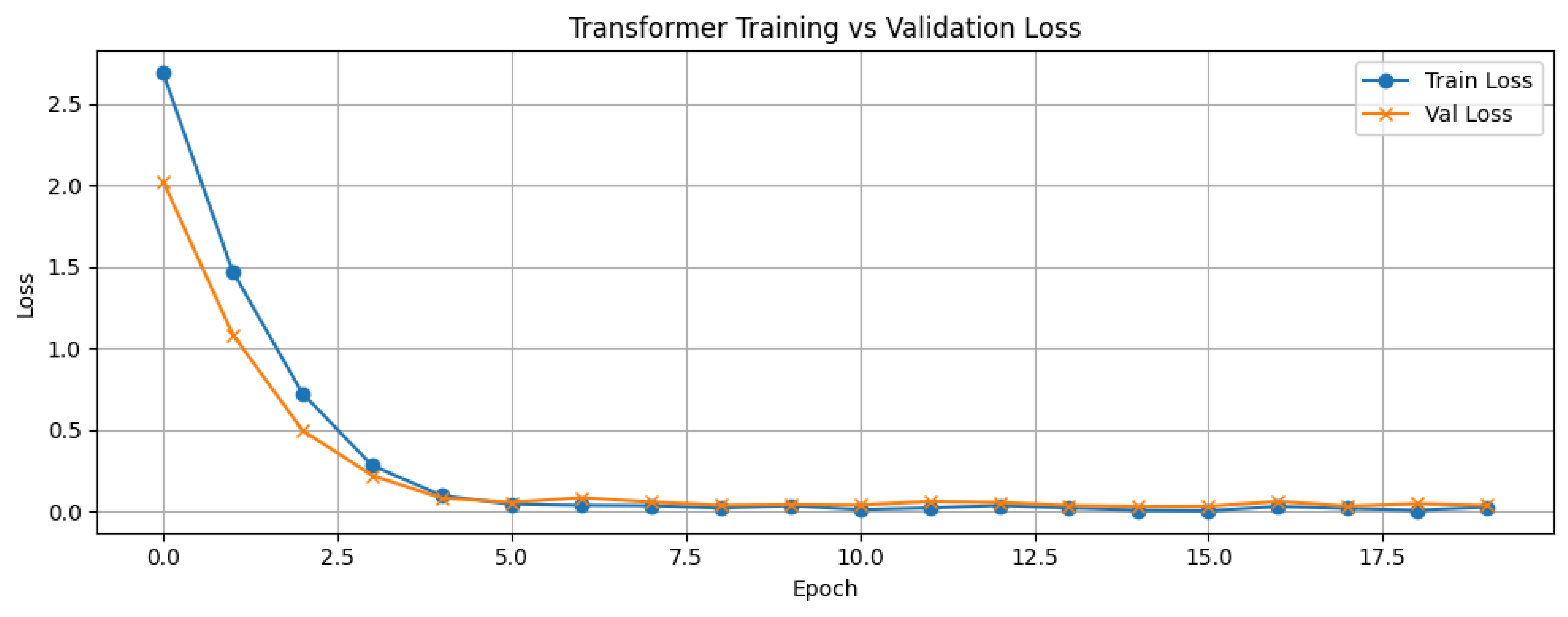
Graphs : Training Curves

- **Encoder-Decoder (LSTM) Without Attention (Luong)**

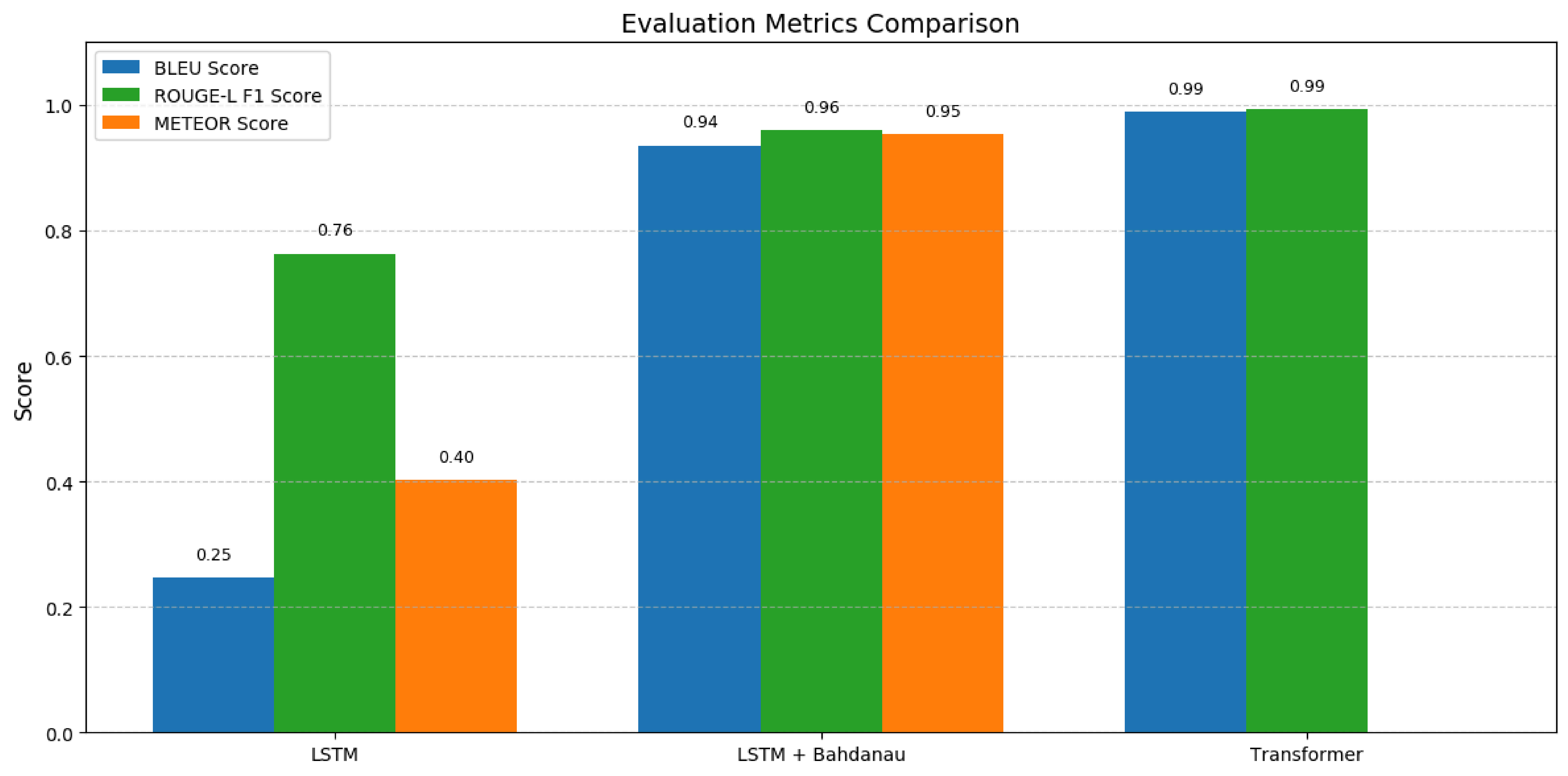


Graphs : Training Curves

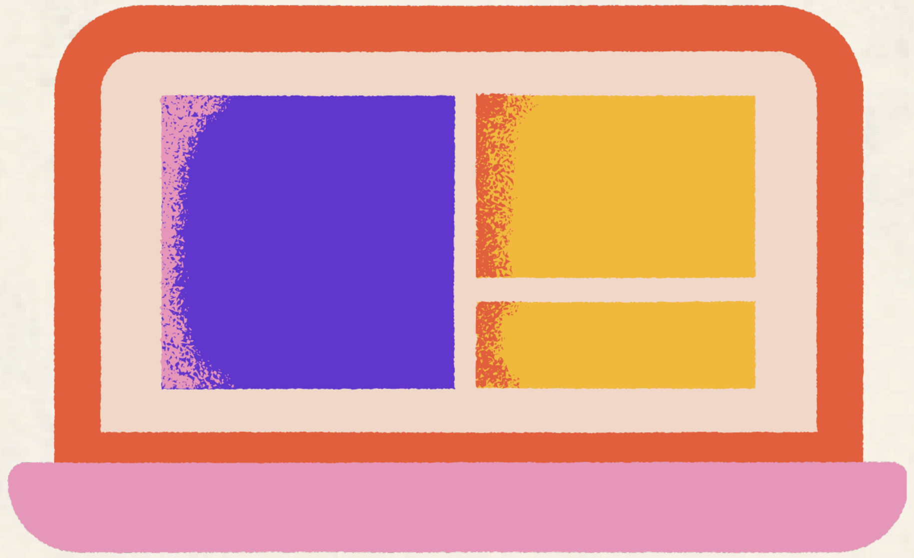
- **Encoder-Decoder (Transformer) Self Attention**



Final Analysis



Conclusion



- Attention mechanisms, particularly Luong Attention, improve grammar correction models over basic LSTM architectures.
- They allow the decoder to focus on relevant parts of the input sentence. Benefits include:
-
    - Higher accuracy
    - Lower training loss
    - More fluent and contextually relevant paraphrases
- Overall, attention-based models significantly enhance the performance of grammar correction tasks.

Thank You

Canva