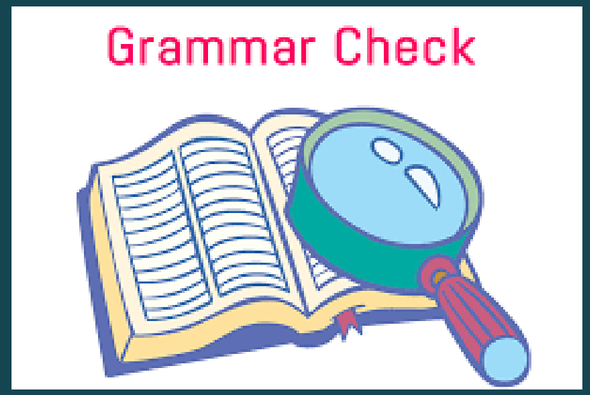
MIT | Academy of Engineering

DEEP LEARNING PRACTICAL EXAMINATION



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

SCHOOL OF COMPUTER ENGINEERING

Our Team

Project Guide:

Mrs.Diptee Ghusse

Students Name:

- Shruti Deokar (202201040207)
- Riya Garasangi (202302040014)

INDEX

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. Models are trained on sentence pairs and assessed using BLEU and ROUGE scores, showing that attention significantly improves performance.

INTRODUCTION

Paper Summary

- **Paper Title**: An Automatic Grammar Error Correction Model Based on Encoder-Decoder Structure for English Texts
- **Authors:** Jiahao Wang, Guimin Huang*, Yabing Wang
- **Year:** 2022
- **Publication:** EAI Endorsed Transactions on Scalable Information Systems
- **Observations**
- The paper proposes a dual-encoder model (DCIM) for English grammar error correction (GEC), combining a Transformer-based context encoder and a Bi-GRU encoder to capture both local and global sentence features.
- Achieved state-of-the-art performance on benchmarks (CoNLL-2014, JFLEG), with a 70.7% precision and 60.2% F0.5 score, outperforming baseline models like Nested-GRU and MLConv.

Paper Summary

- **Challenges**
- Resource limitations: Performance gaps compared to large-scale ensemble models (e.g., Grundkiewicz et al.) due to smaller parameter size and training data.
- Error diversity: Handling rare or nuanced grammar errors requires further model refinement.
- Scalability: The model's lightweight design trades off some accuracy for practicality, suggesting future work on parameter compression techniques.

PROBLEM STATEMENT

Manual grammar correction is time-consuming and error-prone. There is a need for an automated system that can accurately correct grammatical errors in English text using deep learning techniques.

OBJECTIVES TO BE ACHIEVED

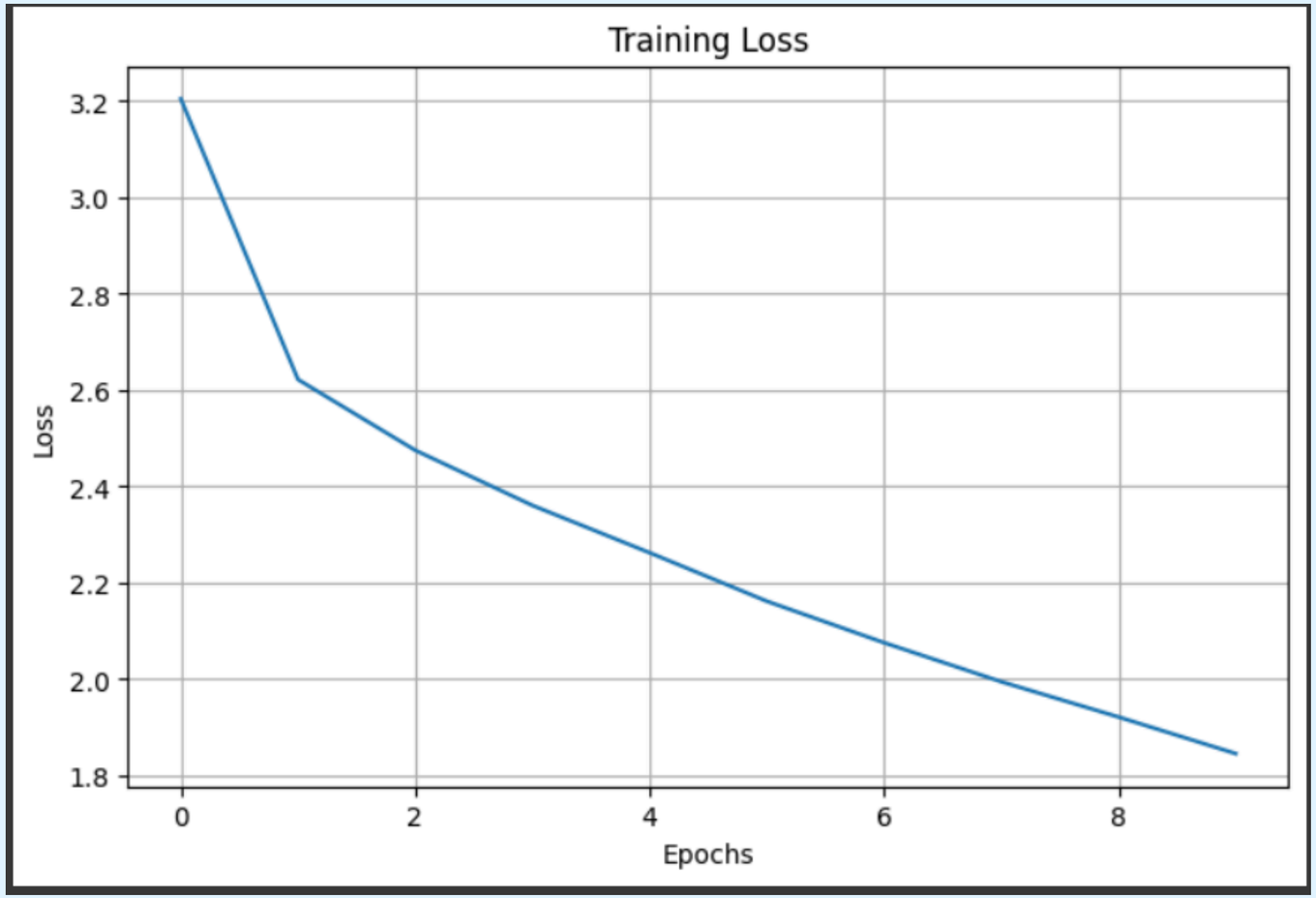
- To implement three encoder-decoder architectures for grammar correction:
- LSTM without attention
- LSTM with Bahdanau attention
- Transformer with self-attention
- To evaluate and compare their performance using standard metrics.
- To analyze the impact of attention mechanisms on model accuracy and efficiency.

Dataset Details

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

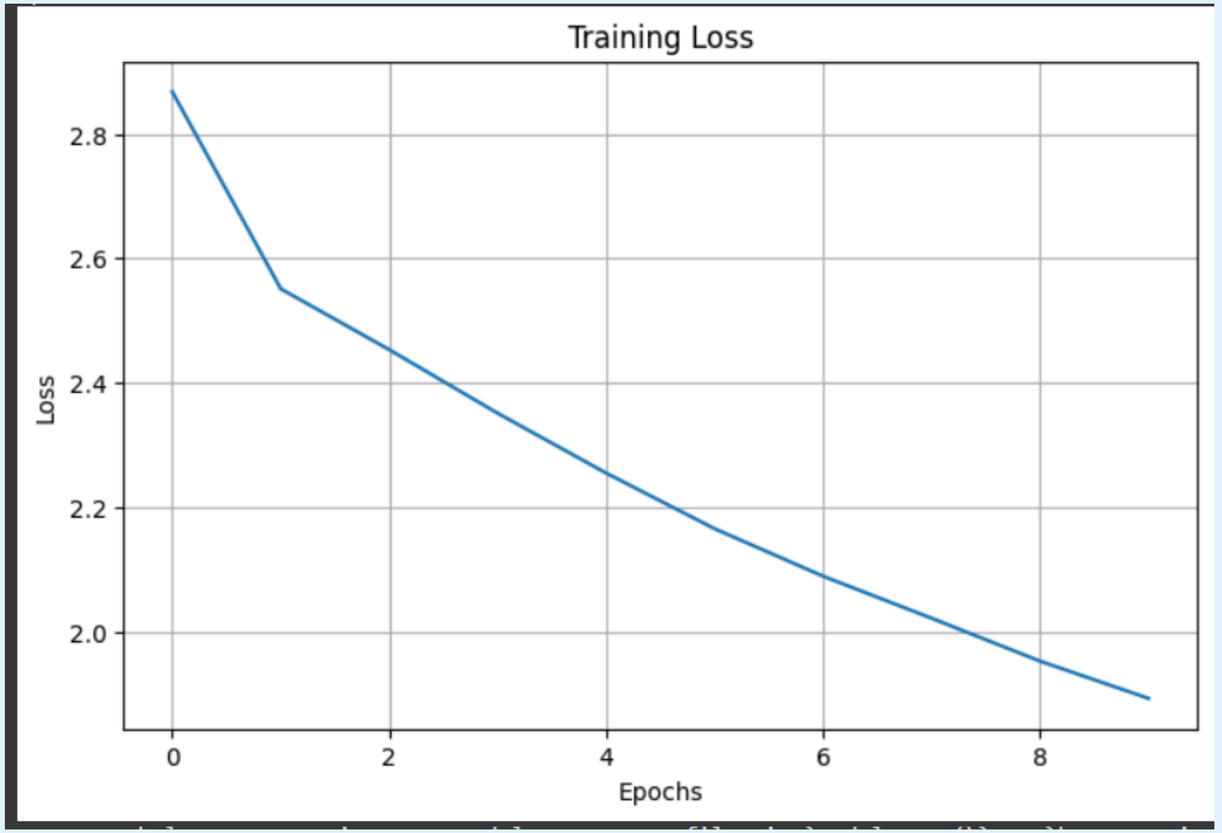
Graph Analysis

LSTM Without Attention



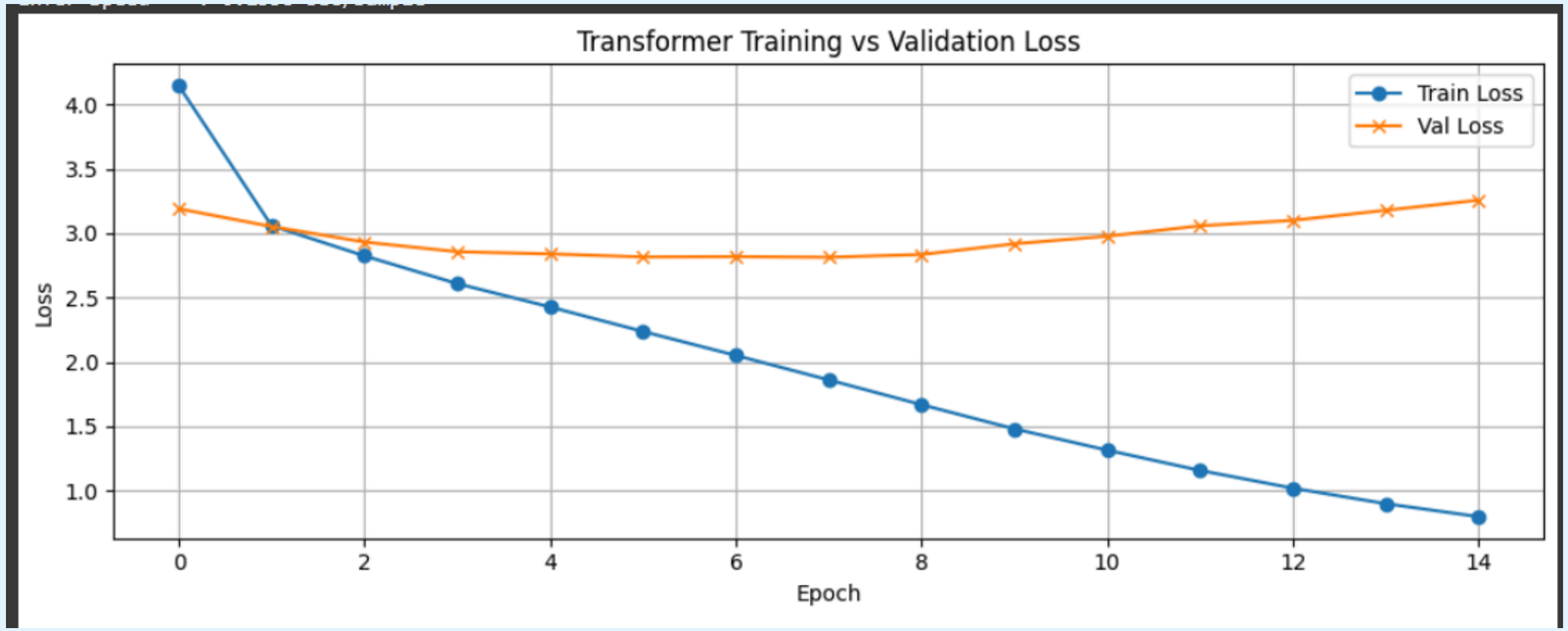
Graph Analysis

LSTM With Attention (Bahdanau)

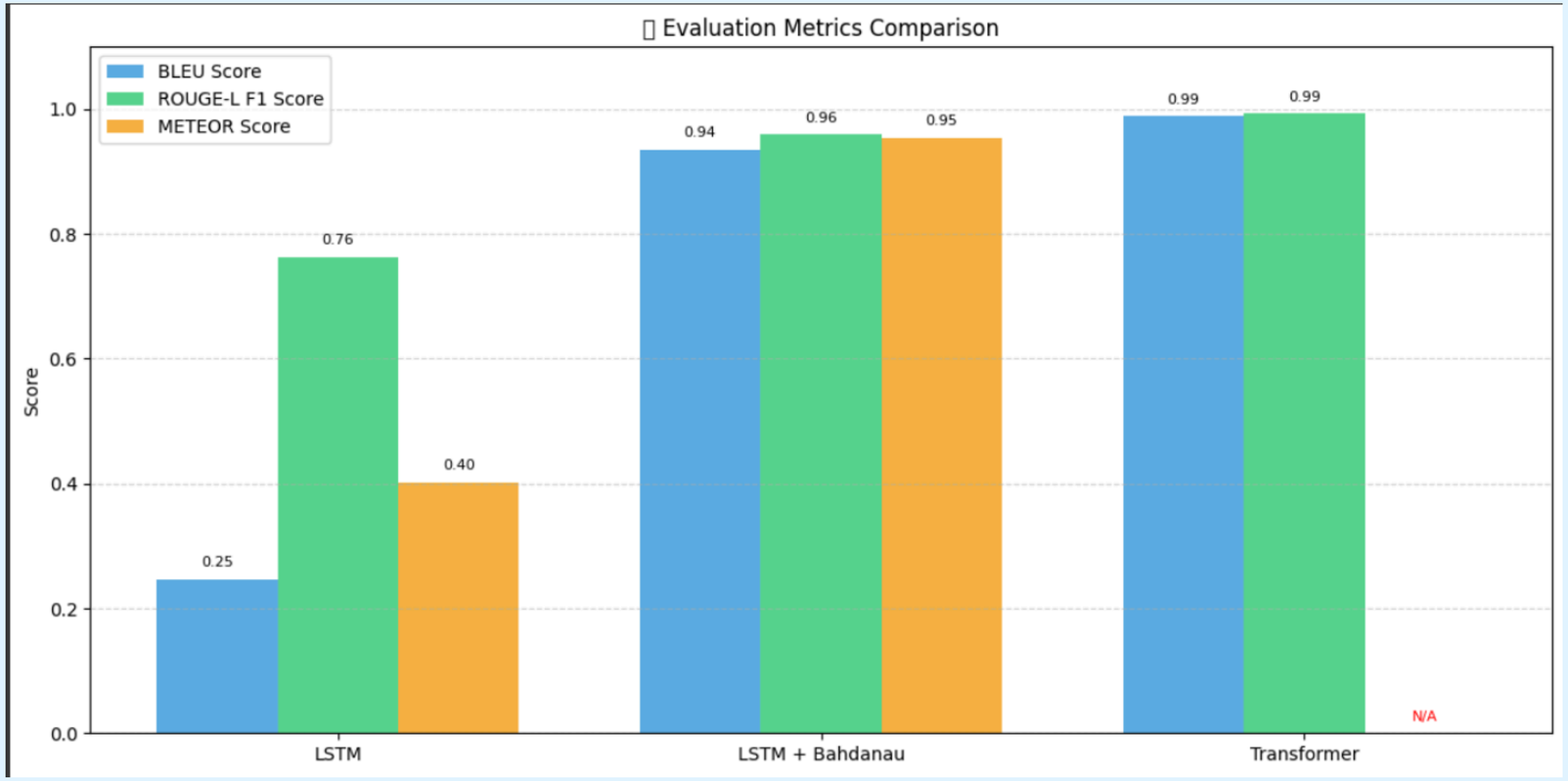


Graph Analysis

Self Attention (Transformer)



Models Performance Analysis



Thank you!