MIT Academy of Engineering
(An Autonomous Institute)

English Grammar Correction

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

Group Members
Md Yaseen Alam
Vaibhav Jadhav
Om Borle
Vidya Bingi

Contents

RESEARCH PAPER SUMMARY

1. ## AIM :

To build an accurate grammar correction system using deep learning.

2. ## OBJECTIVES:

- Identify limits of traditional grammar tools.
- Develop a deep learning-based correction model.
- Use seq2seq with attention for better context.
- Evaluate model performance on benchmark datasets.

3. ## PROBLEM STATEMENT :

Traditional grammar tools rely on fixed rules and lack context understanding, leading to low accuracy. A data-driven deep learning approach is needed for better correction.

METHODOLOGY

1. **Model Architecture:**
   - Utilized a Transformer-based encoder-decoder structure.
   - Leveraged pre-trained BERT for the encoder to benefit from contextual embeddings.
   - Implemented a Transformer decoder initialized with random weights.
2. **Tokenization & Preprocessing:**
   - Applied WordPiece tokenization for input and output sequences.
   - Introduced special tokens [CLS], [SEP] for BERT input formatting.
3. **Input Representation:**
   - Input sentences passed to BERT to generate contextualized token embeddings.
   - Decoder receives shifted target sentences for training.

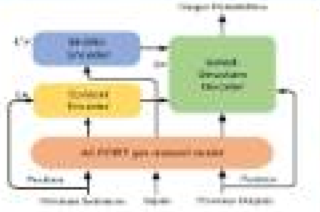


Figure 1. The structure of DCBM

METHODOLOGY

4.**Training Strategy:**
- Used teacher forcing to guide decoder training.
- Employed cross-entropy loss as the optimization objective.

5.**Data Used:**
- Trained and evaluated on public grammar correction datasets (e.g., CoNLL-2014, JFLEG).

6.**Evaluation Metrics:**
- Evaluated using Precision, Recall, $F_{0.5}$-score, and GLEU score for effectiveness.

7.**Inference:**
- Applied beam search during inference to generate corrected sentences.

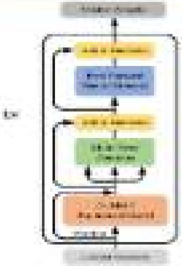
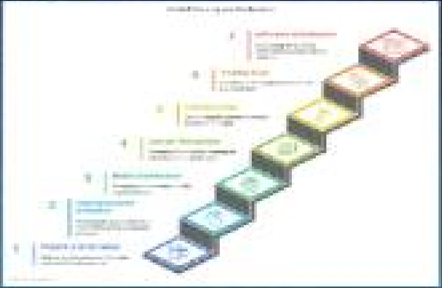


Figure 3. The structure of context information encoder

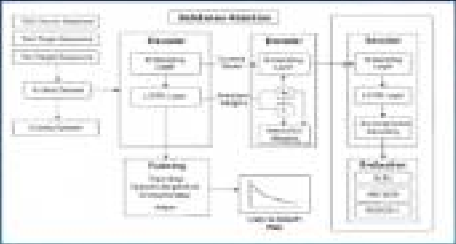
MODEL DIAGRAM & ARCHITECTURE

- Without Attention:



MODEL DIAGRAM & ARCHITECTURE

- With Attention:



MODEL DIAGRAM & ARCHITECTURE

- **Self Attention:**

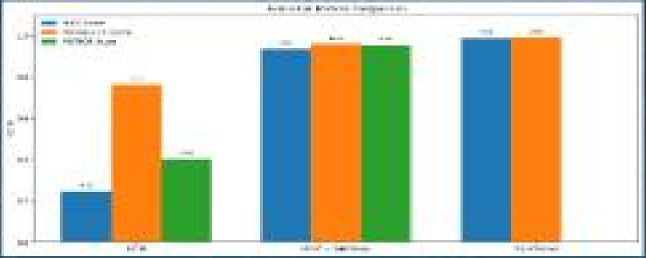
DATASET DESCRIPTION

- Dataset Name: Grammar Correction Dataset
- Format: CSV (Comma-Separated Values)
- Purpose: Created for training and evaluating grammar correction models in Natural Language Processing (NLP).
- Total Records: [illegible]
- Columns:
- **Error Type** – Category of the grammatical mistake.
- **Ungrammatical Statement** – The incorrect sentence containing grammar issues.
- **Standard English** – The corrected version of the sentence.

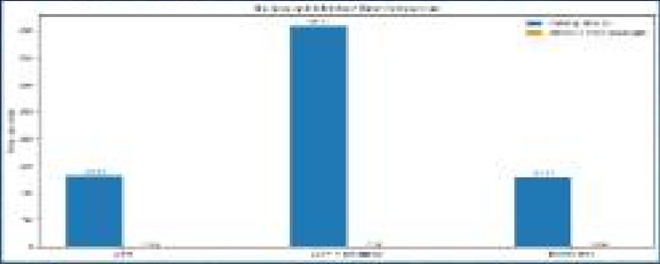
Use Cases:
- Grammar correction models (NLP)
- Grammar error detection and correction tools
- Language learning applications

[illegible] For:
- Training supervised machine learning models for sentence correction.
- Fine-tuning transformer-based models like BERT or T5.

EVALUATION METRICS COMPARISION :

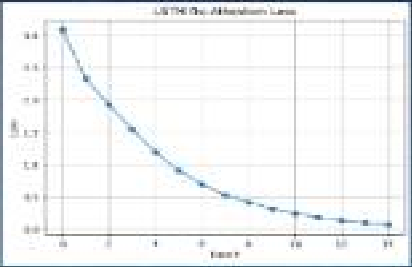


TRAINING AND INFERENCE TIME COMPARISION :



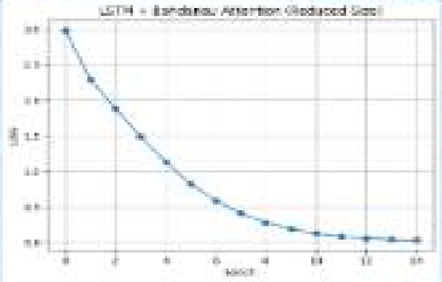
GRAPHS

1. WITHOUT ATTENTION



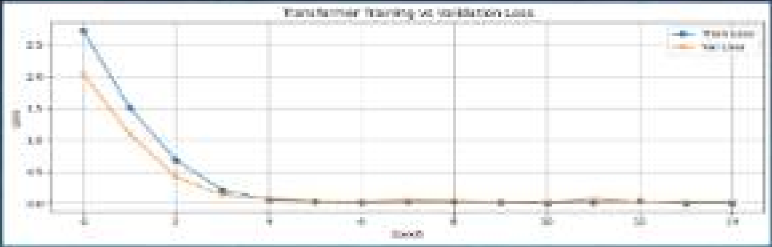
GRAPHS

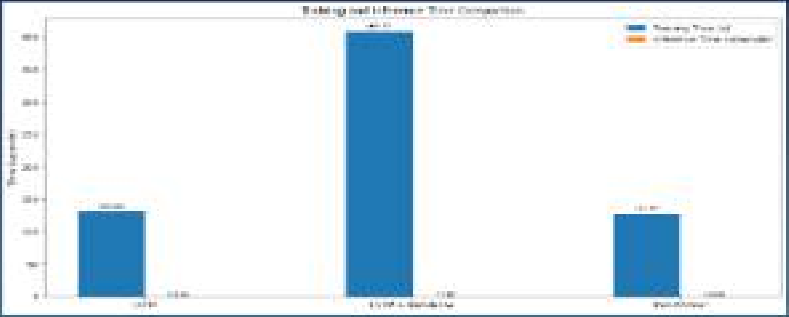
1. WITH ATTENTION

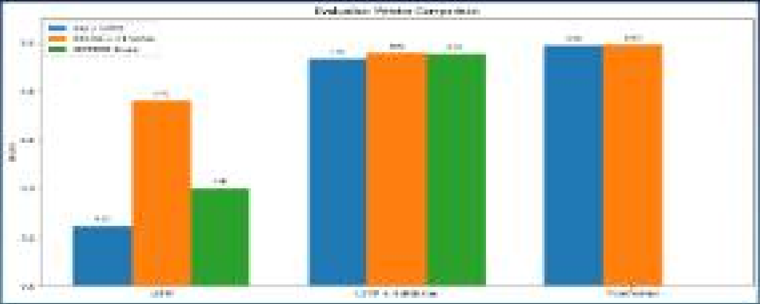


GRAPHS

1. SELF ATTENTION







CONCLUSION

The comparative study demonstrates that the Transformer model, with its self-attention mechanism, significantly outperforms both LSTM and Bahdanau attention models in terms of grammer correction accuracy and processing efficiency. Its superior evaluation scores and faster inference time make it the most suitable choice for real-time grammar correction applications.


THANK
YOU!