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

Dialogue Generation

Using Encoder-Decoder Architectures with Attention Mechanisms

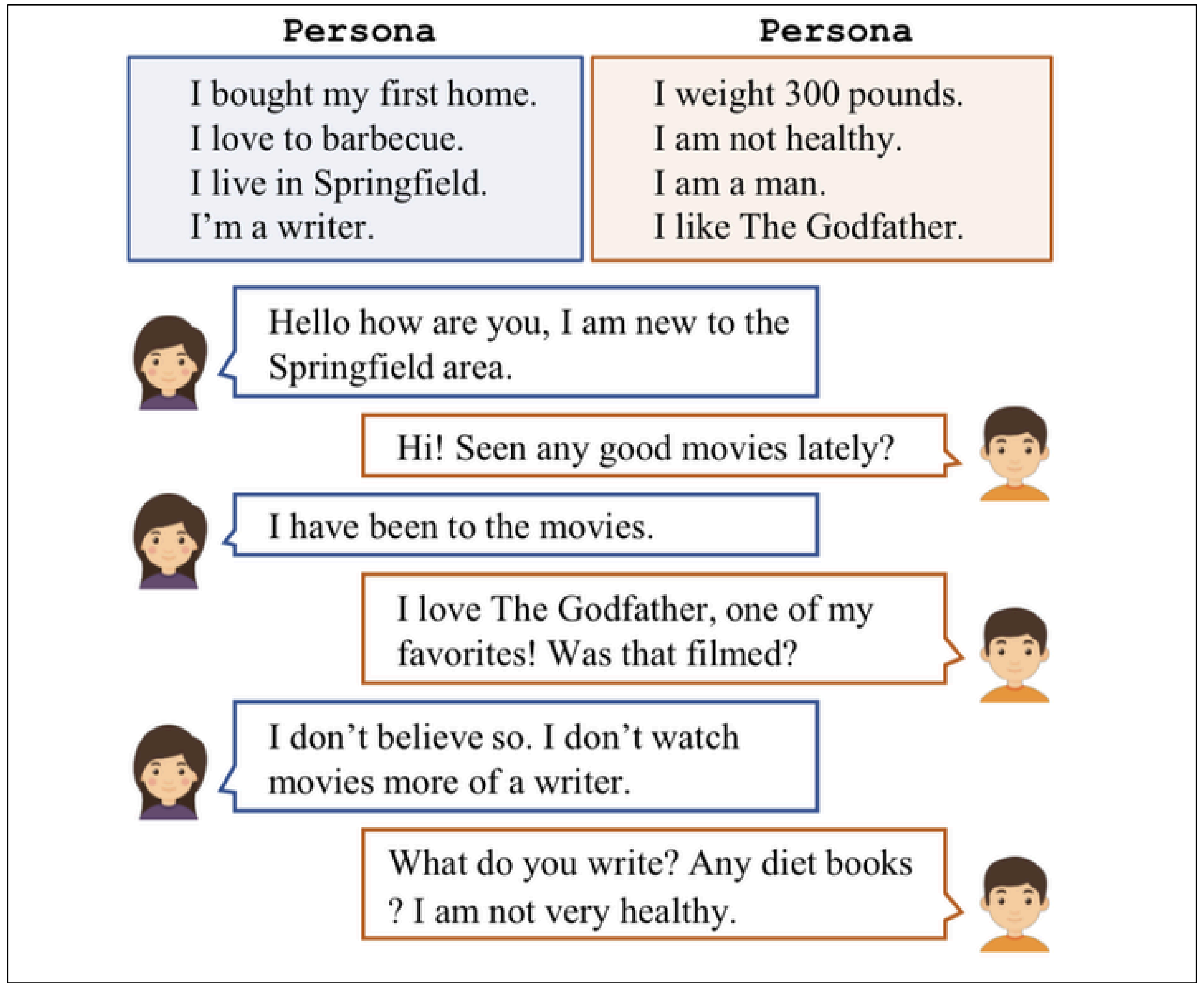
Members:
Shravani Pidurkar (202201070027)
Janhavi Pawar (202201070035)
Tanvi Tonge (202201070053)

Guide: Sunita Barve

Content

Introduction

- Dialogue generation is a deep learning task that generates coherent and contextually relevant responses to user input in a conversation.
- Aim: To implement and compare different encoder-decoder architectures for Dialogue Generation.
- We explore the following models:
  - Without Attention: LSTM-based encoder-decoder
  - With Attention: Luong Attention mechanism
  - With Self-Attention: Transformer architecture



Paper Summary

Aim : Improve dialogue coherence by modeling speaker roles in multi-turn conversations.

Objectives :

- Speaker-aware modeling to distinguish between queries (Speaker-Q) and responses (Speaker-R).
- Parallel hierarchical encoder-decoder to separately process speaker-specific context

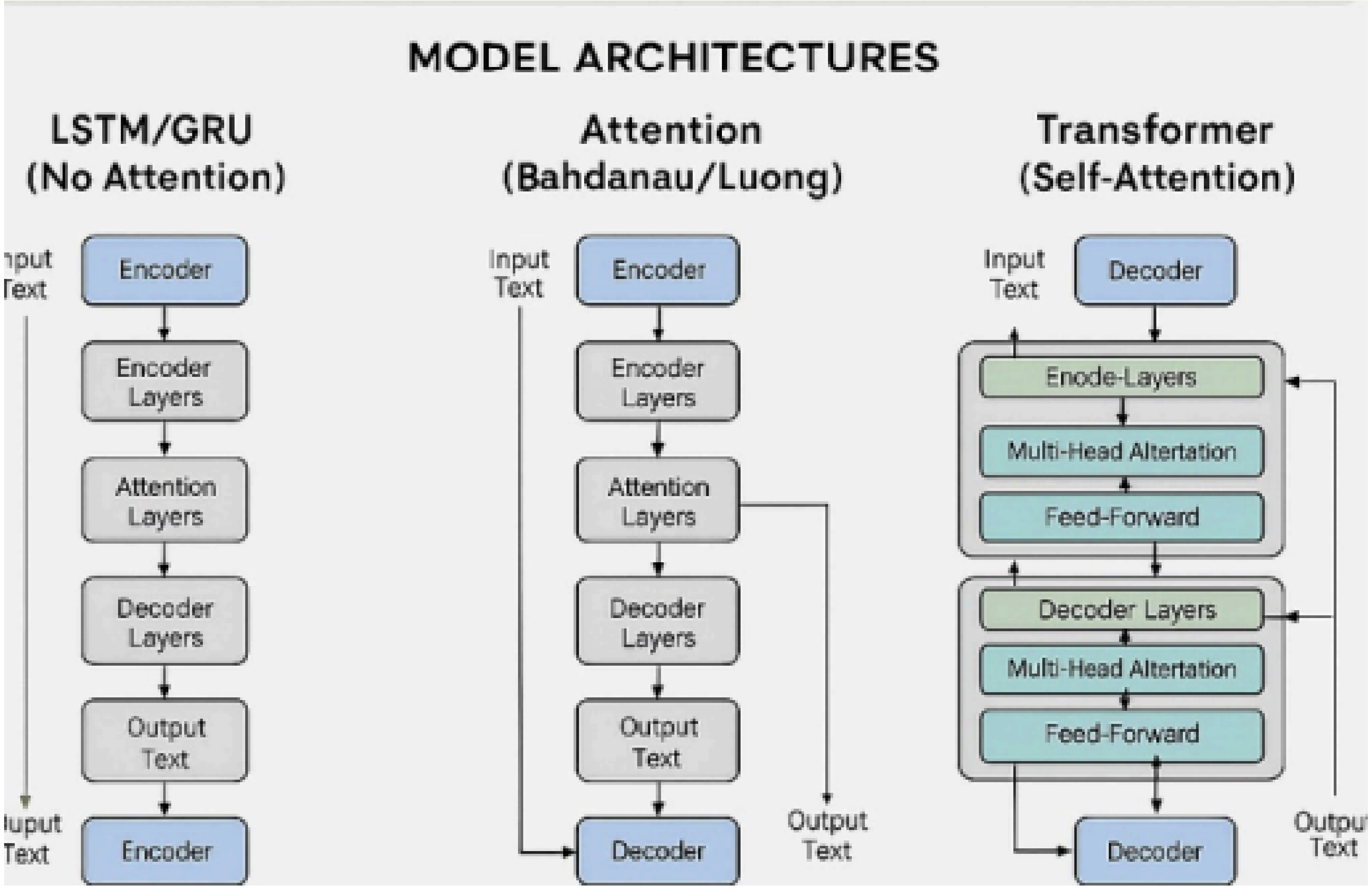
Problem Statement :

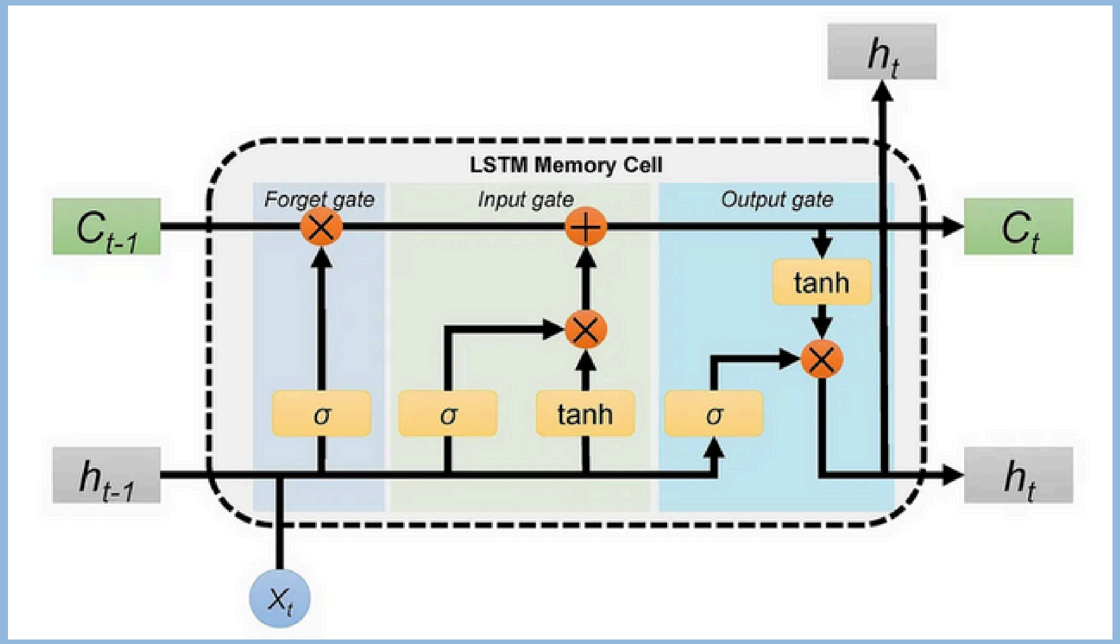
Existing models ignore speaker identity, leading to inconsistent responses. Traditional approaches treat dialogue history as a single sequence, missing speaker-level context.

Methodology

1. **Dataset Preparation**
    - Use a multi-turn dialogue dataset
    - Preprocess: tokenize, clean text, and convert to sequences.
2. **Model Implementation**
    - Seq2Seq with LSTM/GRU (no attention) as baseline.
    - Seq2Seq with Bahdanau attention for improved context handling.
    - Transformer model with self-attention for capturing long-range dependencies.
3. **Training**
    - Use teacher forcing and cross-entropy loss.
    - Apply padding and masking for sequence handling.
4. **Evaluation**
    - Evaluate models using BLEU, METEOR, and ROUGE scores.
5. **Comparison**
    - Analyze performance to understand the effect of attention on dialogue quality and coherence.

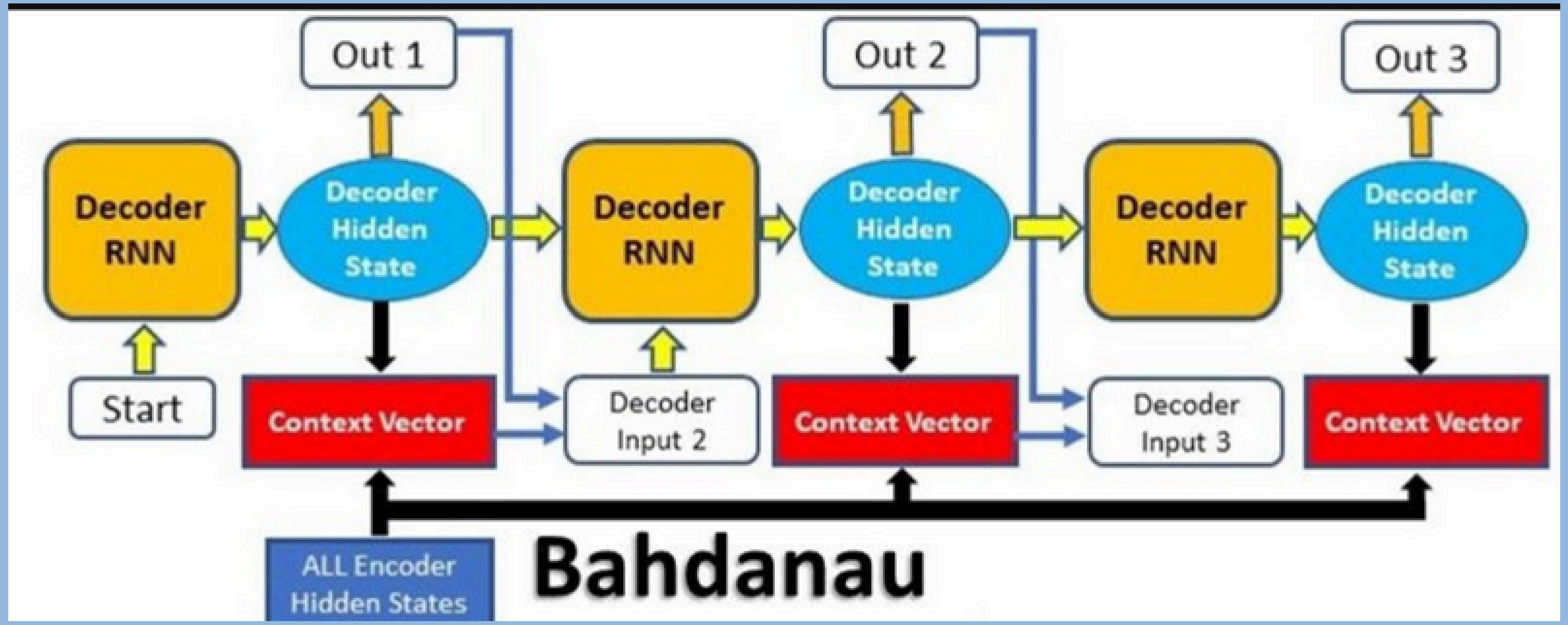
Model Architecture

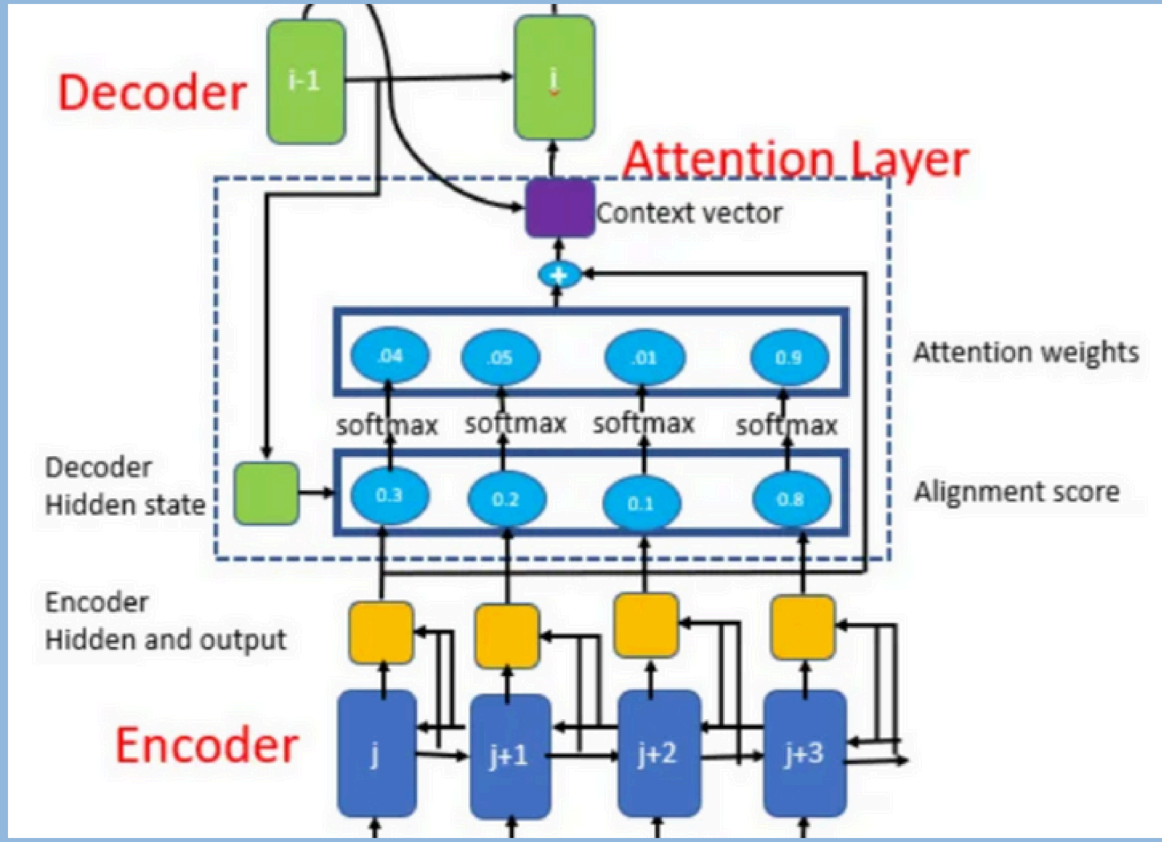




Without Attention LSTM-based architecture

With Attention Bahdanau artitecture



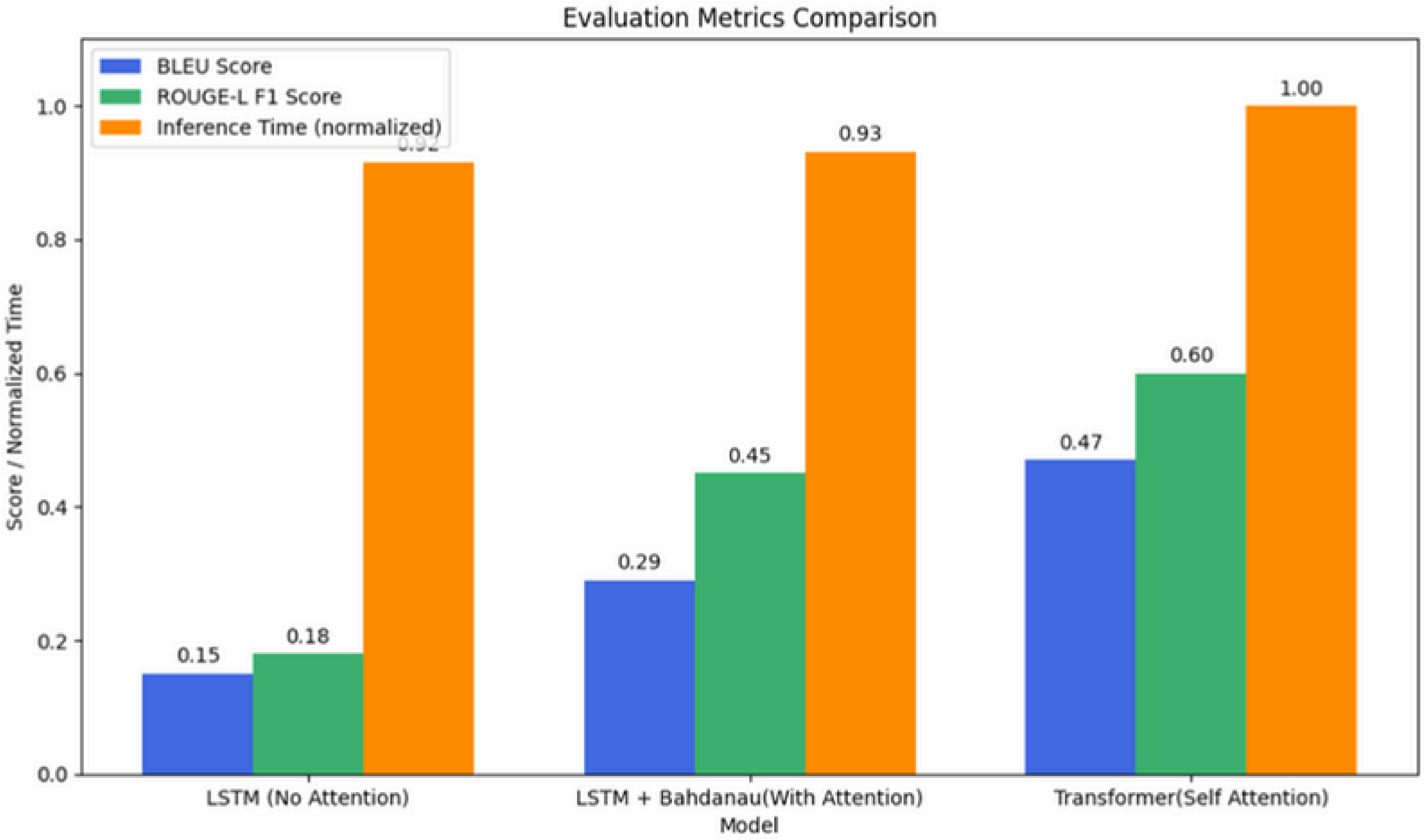


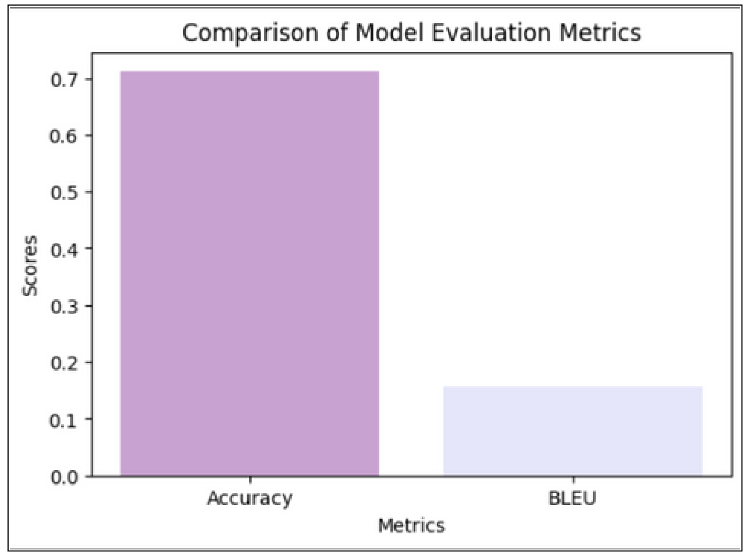
With Self-Attention Transformer Artitecture

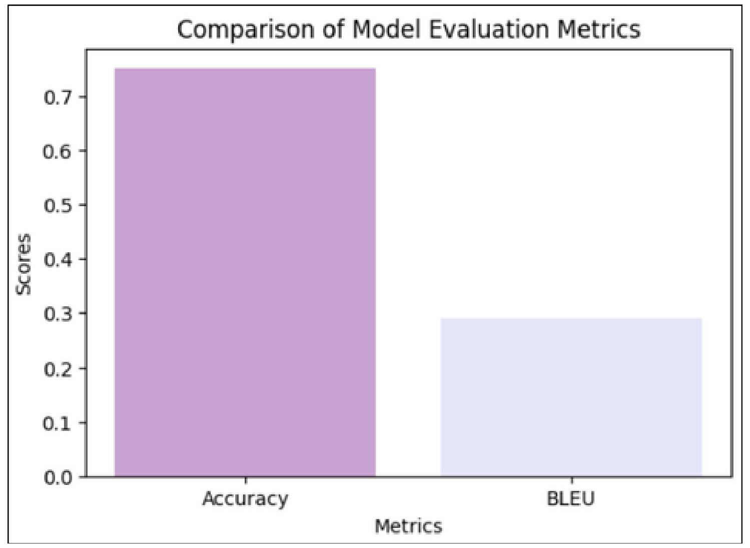
Dataset Description

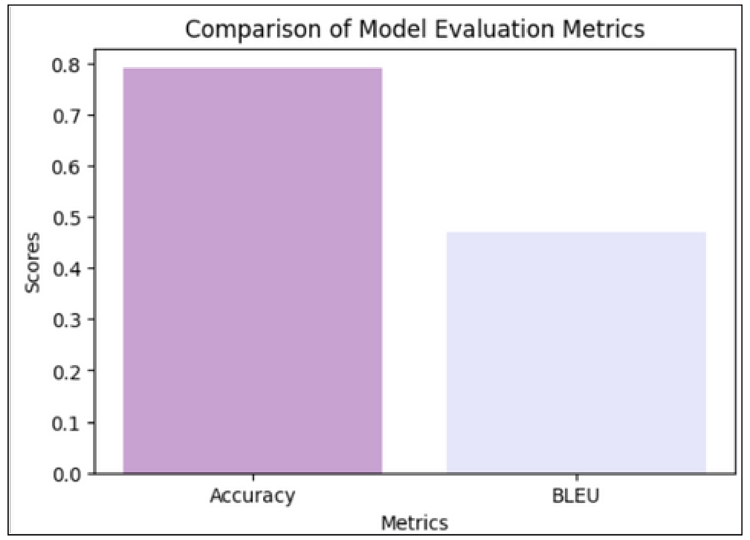
- This dataset is used for research or training of natural language processing (NLP) models. The dataset include various types of conversations such as casual or formal discussions, interviews, customer service interactions, or social media conversations.
- Source : Kaggle
- Dataset Size : 3510 unique values

Evaluation and Results









Results

The LSTM model performed reasonably well in detecting anomalies, benefiting from its ability to capture sequential patterns. However, it exhibited slight overfitting—indicating strong performance on training data but reduced generalization to unseen data—and struggled with long-range dependencies, a known limitation of traditional RNNs. Incorporating Bahdanau Attention helped mitigate these issues by allowing the model to focus on relevant parts of the input sequence dynamically, improving both convergence speed and accuracy. This mechanism enabled the model to better align with important input features during anomaly detection, achieving the highest overall performance. On the other hand, the Transformer model, which relies entirely on self-attention, generalized well and effectively captured global context across sequences, but required more training time and computational resources. Overall, the use of attention mechanisms significantly enhanced both the interpretability and effectiveness of the models in detecting anomalies.

Conclusion

This project demonstrated the implementation and comparison of various encoder-decoder architectures for dialogue generation. Models without attention generated basic but often generic responses, while attention-based models (Bahdanau) improved the contextual relevance and coherence of replies. Transformer-based models further enhanced performance by using self-attention to capture complex dependencies in conversational context. Evaluation results confirmed that attention mechanisms significantly improve the quality and fluency of generated dialogues. Overall, the project highlights the potential of deep learning in building effective and context-aware conversational systems for real-world applications.

Thank You !