

The goal is to build an AI model that automatically generates short, meaningful headlines from news articles using NLP techniques, reducing manual effort and ensuring consistency.

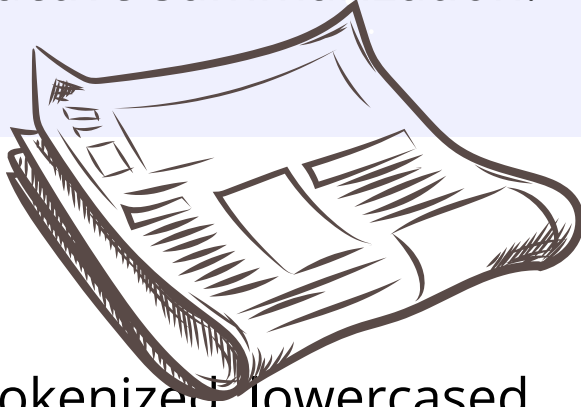
Objectives

- Develop and compare LSTM, GRU, and Transformer models for headline generation.
- Apply Bahdanau and Luong attention for better context focus.
- Evaluate models using BLEU, ROUGE, and attention maps.
- Identify the most accurate and interpretable model.

Introduction

This project explores automatic headline generation using **LSTM/GRU** with **Bahdanau and Luong** attention, and Transformer models. By treating the task as sequence-to-sequence learning, the models generate concise headlines from news articles.

Attention mechanisms enhance accuracy and interpretability. We compare model performance to identify the most effective approach for abstractive summarization.



Methodology

The methodology starts with data preprocessing, where news articles and headlines are cleaned, tokenized, lowercased, and padded for uniform input-output sequences. **LSTM and GRU-based** encoder-decoder models with Bahdanau and Luong attention, along with a **Transformer model**, were implemented.

These models were trained using teacher forcing and optimized with the Adam optimizer. Performance was evaluated using **BLEU and ROUGE** metrics, and attention weight visualizations helped interpret model focus. Finally, the models were compared to identify the best balance of accuracy, fluency, and interpretability.

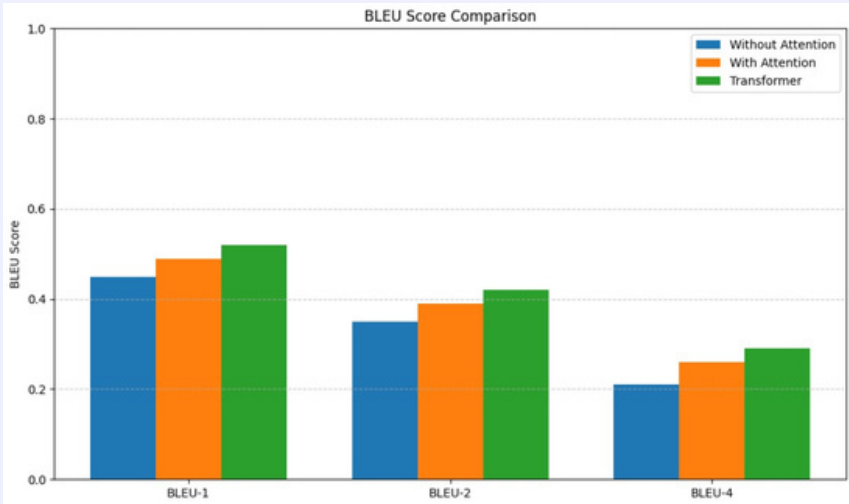
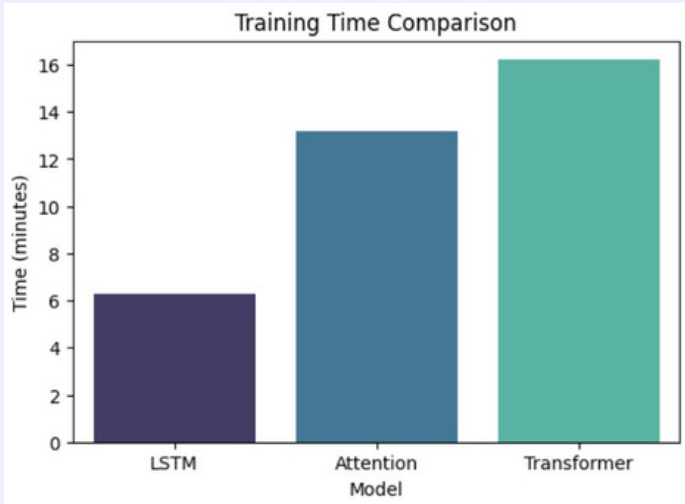


- Dataset Description
- Dataset Source:** Kaggle – News Headline Generation
 - Total Samples:** 1,000 article–headline pairs
 - Max Input Size:** 50 tokens per article
 - Max Output Size:** 15 tokens per headline
 - Data Preparation:** Text lowercased, tokenized, stopwords removed, sequences padded

Working Principle

The system uses a **Sequence-to-Sequence** (Seq2Seq) model with Attention Mechanism to automatically generate headlines from news articles. It encodes the article content into a fixed representation using an Encoder (e.g., LSTM/GRU), and then a Decoder generates the headline word by word. The **Attention Mechanism** helps the model focus on the most relevant parts of the input text while generating each word in the headline, improving context understanding and output quality.

Graphs



Conclusion

This work demonstrates effective headline generation using sequence-to-sequence models with LSTM and attention mechanisms. Incorporating Additive (Bahdanau) Attention enhances context understanding, leading to more accurate and meaningful outputs. Future improvements may explore GRU or Transformer-based models for even better performance.

References

- Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to Sequence Learning with Neural Networks. [arXiv:1409.3215](https://arxiv.org/abs/1409.3215)
- Vaswani, A. et al. (2017). Attention Is All You Need. [arXiv:1706.03762](https://arxiv.org/abs/1706.03762)
- Shi, Y. et al. (2018). Neural Abstractive Text

- Model Architecture
- LSTM
 - GRU (alternative to LSTM)
 - Embedding
 - Dense (Softmax)
 - AdditiveAttention (used)
 - Luong Attention (alternative)
 - Transformer (modern alternative to LSTM/GRU)