INTERACTIVE HEADLINE GENERATION

Presented by -
Sanchit Panda - 202201040041
Vaishnavi Kedar - 202201040045
Ishika Bondre - 202201040057
Sameer Sable - 202201040064

Contents:

INTRODUCTION

This project focuses on automated headline generation using deep learning techniques. It compares the performance of Vanilla LSTM, LSTM with Bahdanau Attention, and LSTM with Self-Attention. The models are trained to generate concise, meaningful headlines from news articles, aiming to enhance summarization accuracy through attention mechanisms.

PAPER SUMMARY

AIM :

To develop a transformer-based model (TD-NHG) that improves news headline generation by:

- Enhancing text feature extraction using masked self-attention.
- Reducing repetition and improving accuracy with advanced decoding strategies.

Objective :

To enhance the accuracy and diversity of news headline generation by addressing the limitations of existing models, particularly the lack of parallel processing capabilities and the tendency to produce repetitive or inaccurate headlines.

PAPER SUMMARY

Problem Statement :

Existing news headline generation models, particularly those based on sequence-to-sequence architectures or recurrent neural networks, face challenges:

- Limited parallel processing capabilities.
- Tendency to repeat words, leading to less informative headlines.
- Difficulty in selecting and emphasizing important information from news content.

PAPER SUMMARY

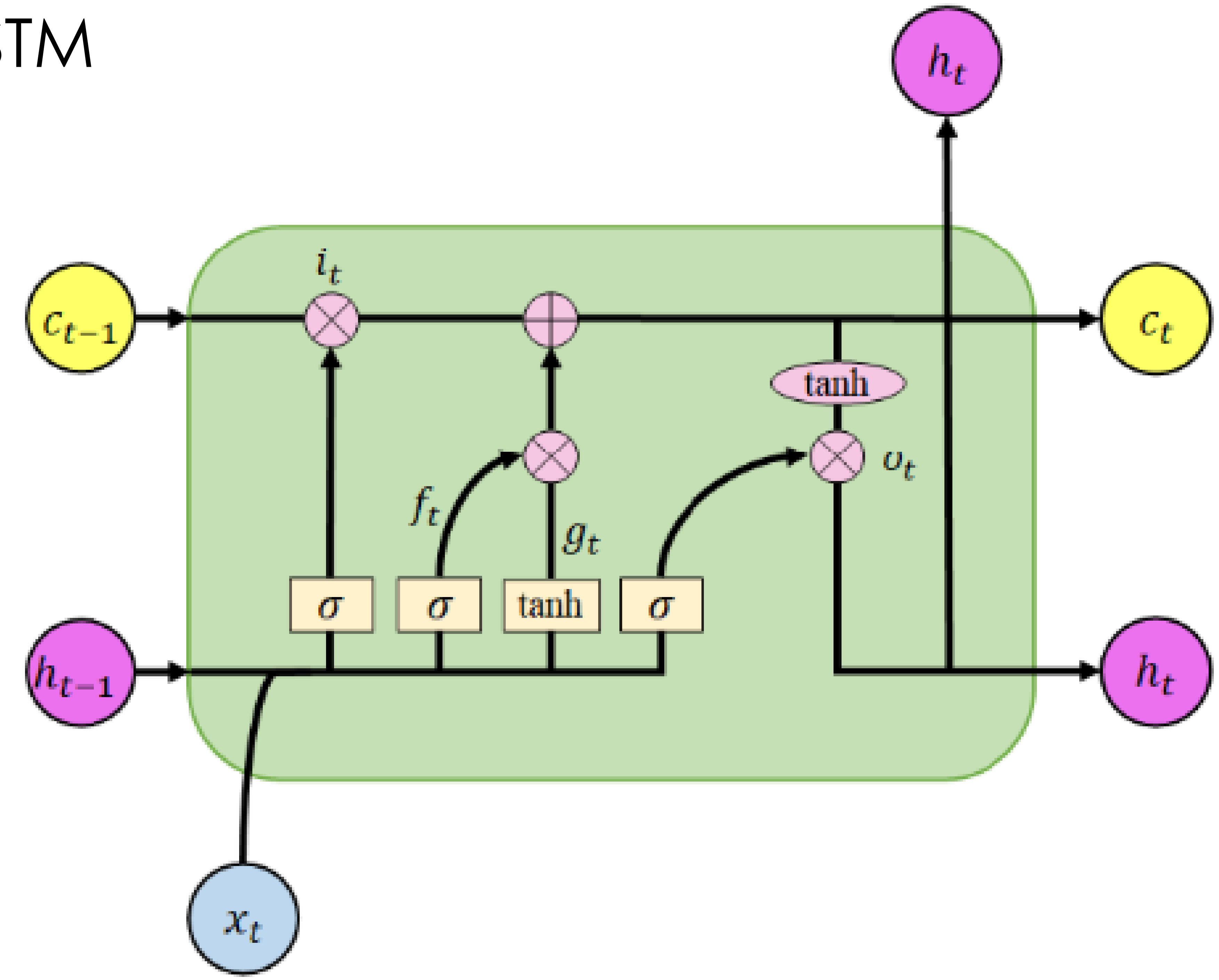
Methodology :

The proposed model, termed TD-NHG (Transformer Decoder for News Headline Generation), incorporates:

- **Masked Multi-Head Self-Attention:** Enables the model to capture feature information from different representation subspaces of news texts.
- **Decoding Selection Strategies:**
  - **Top-k Sampling:** Limits the sampling pool to the top k probable words.
  - **Top-p (Nucleus) Sampling:** Considers the smallest set of words whose cumulative probability exceeds a threshold p.
  - **Repetition Penalty Mechanism:** Discourages the model from generating repetitive words.

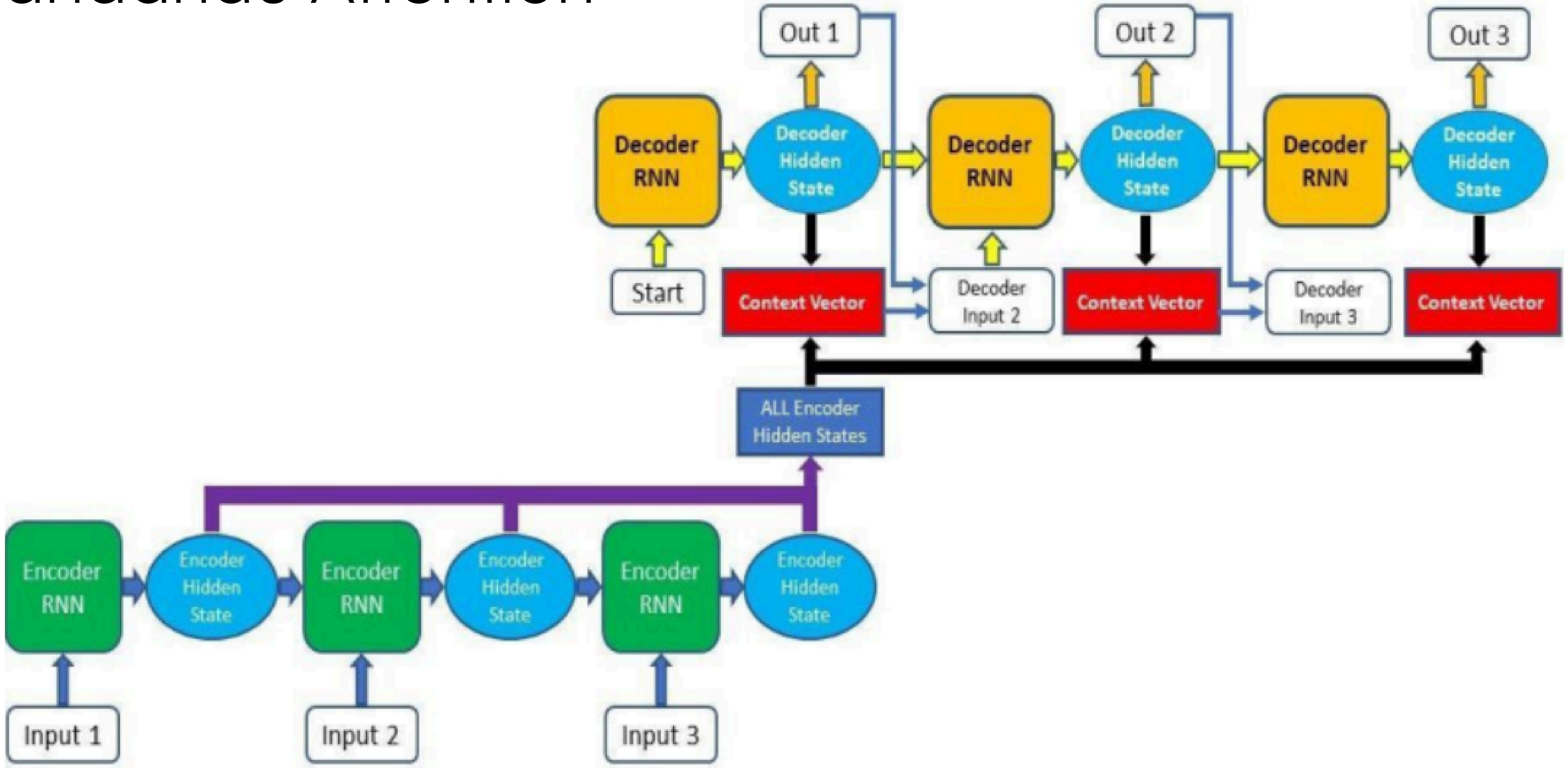
MODEL DIAGRAM AND ARCHITECTURE

Vanilla LSTM



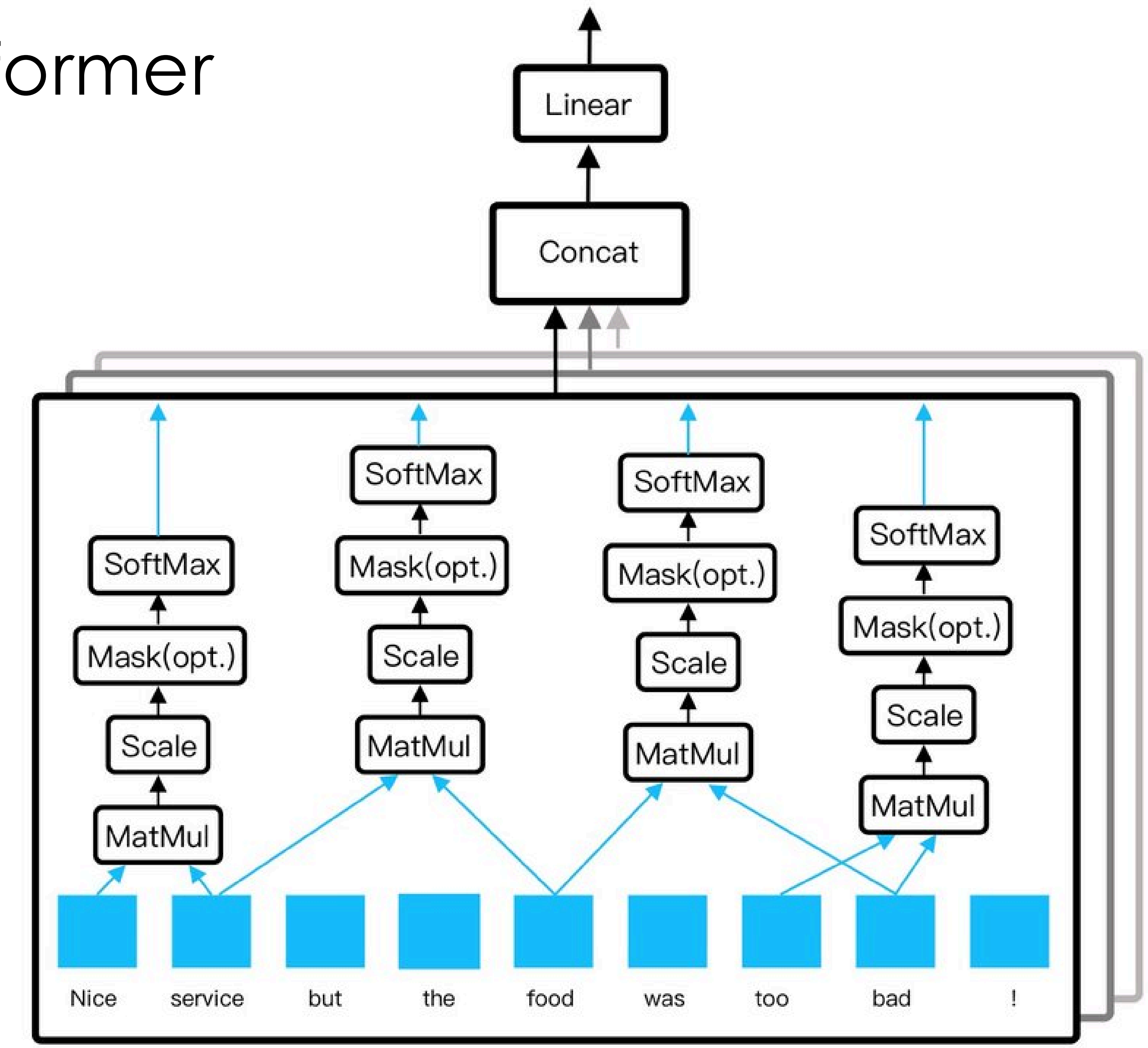
MODEL DIAGRAM AND ARCHITECTURE

LSTM with Bahdanau Attention



MODEL DIAGRAM AND ARCHITECTURE

Transformer



DATASET DESCRIPTION

Dataset Name: **News Summary Dataset**
Date of Data Collection: 03 August 2017
Number of Records (Visible): Approximately 25 entries

Overview:
The News Summary Dataset contains metadata and summarized content of news articles from various Indian and international media sources. It is structured to support natural language processing (NLP) tasks such as text summarization, classification, and content retrieval.

DATASET DESCRIPTION

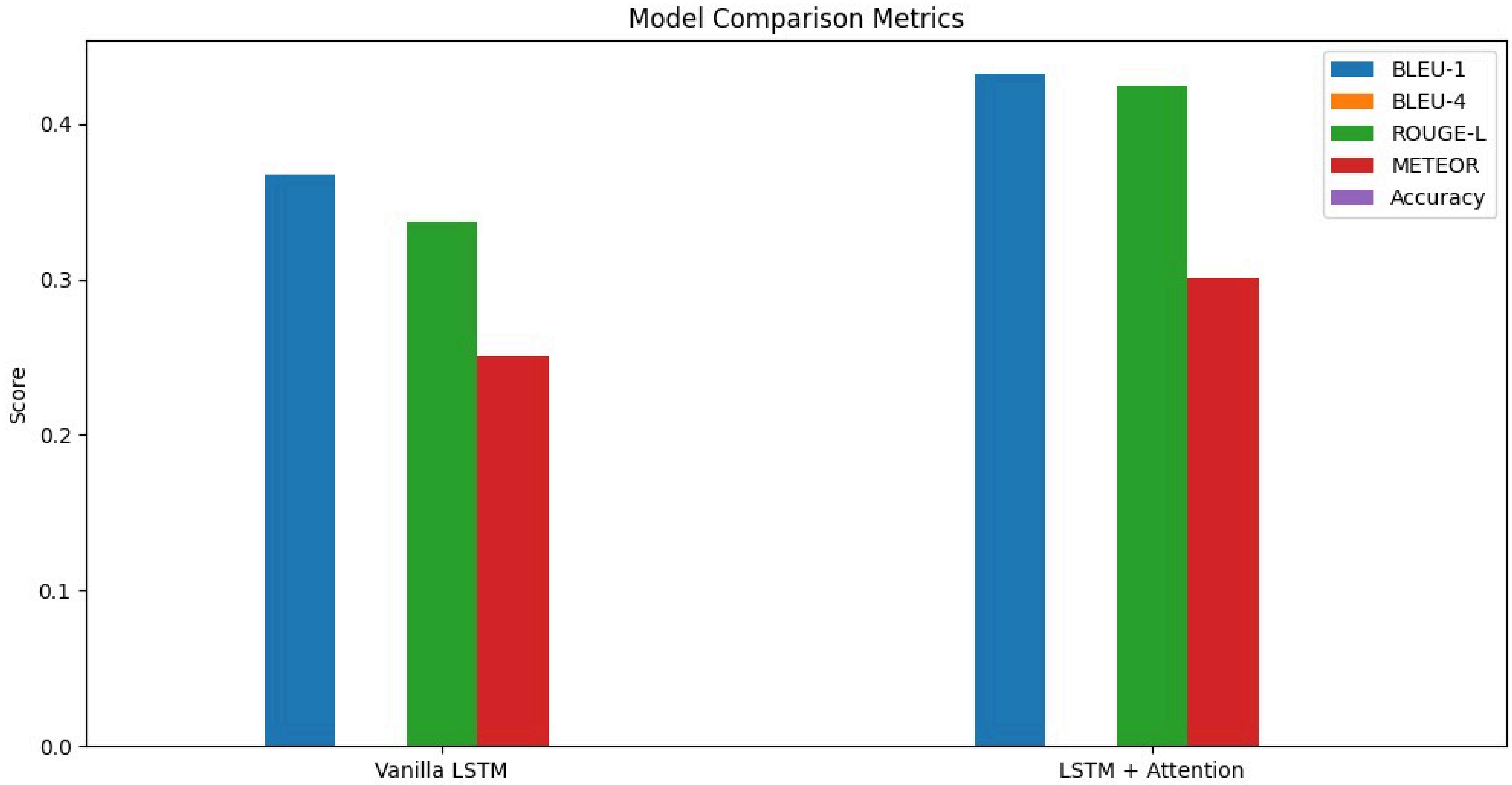
Column	Description
author	The name of the journalist or contributor responsible for writing the article.
date	The date on which the article was published, formatted as "DD MMM YYYY,Day".
headlines	A concise title or headline of the article summarizing the main point.
read_more	A hyperlink to the original full-length article.
text	A brief summary of the article content, suitable for short-text analysis.
ctext	A more detailed summary or full version of the article's content.

DATASET DESCRIPTION

Potential Applications:

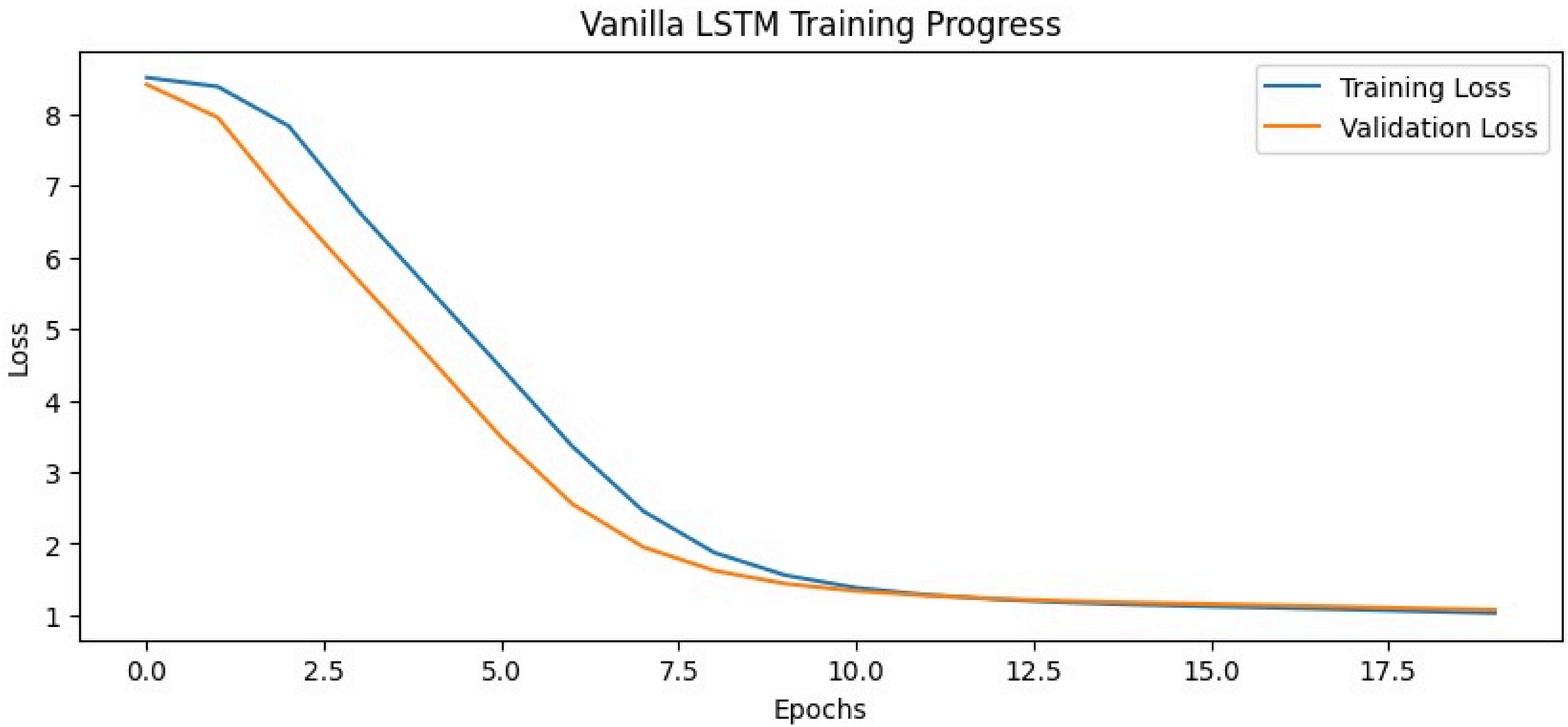
- Development and evaluation of text summarization algorithms (e.g., comparing short vs. long summaries).
- Training datasets for headline generation models.
- Information retrieval and keyword-based content classification.
- Author-wise trend and content analysis.

METRIC-WISE PERFORMANCE COMPARISON



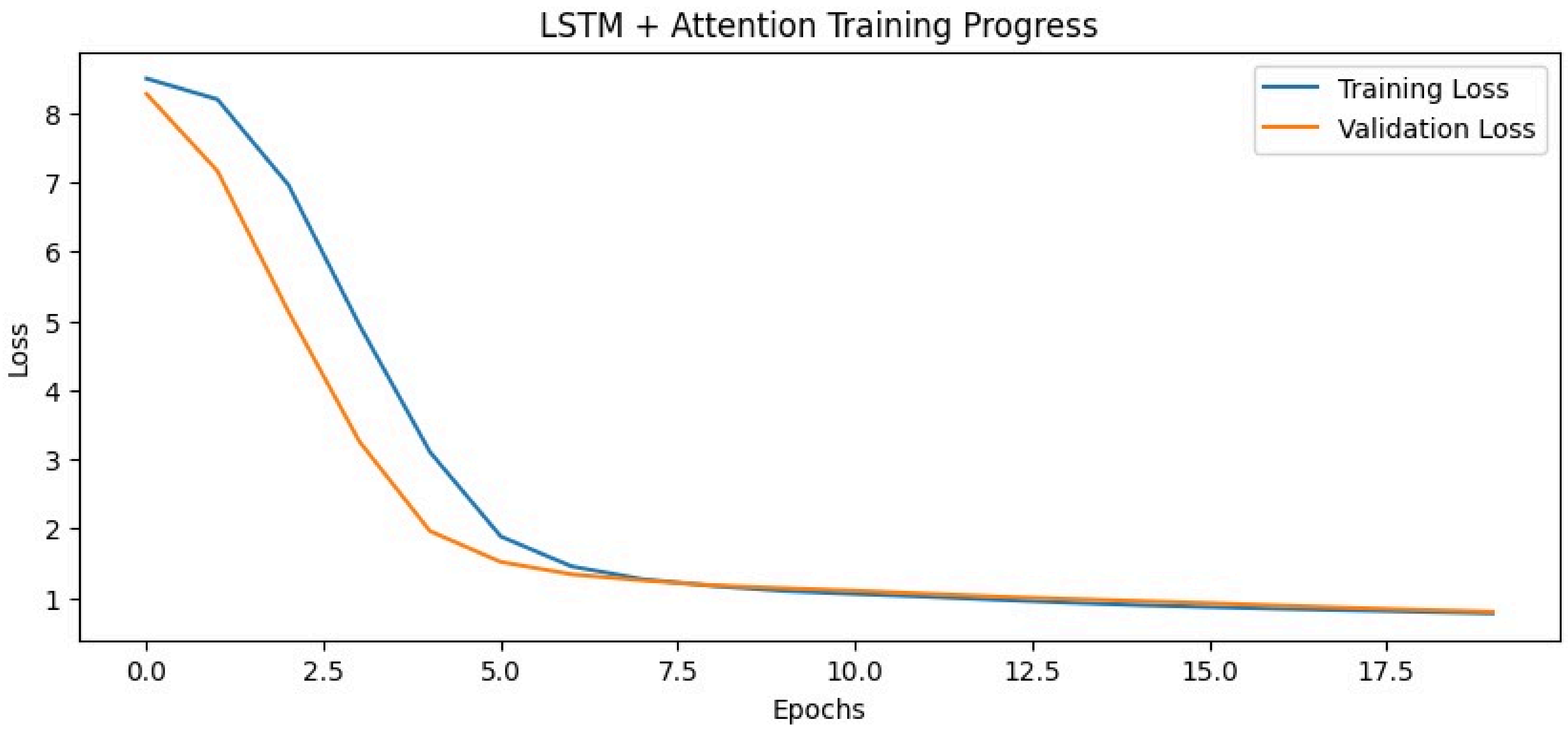
GRAPHS

Vanilla LSTM



GRAPHS

LSTM with Bahdanau Attention



ANALYSIS

Criteria	**Vanilla LSTM** (No Attention)	**LSTM + Bahdanau Attention**	**Transformer (Self-At**
Accuracy / BLEU-4	Low (~0.2)	Medium (~0.35)	High (~0.5)
ROUGE-L F1	~0.25	~0.4	~0.55
METEOR	~0.15	~0.3	~0.45
Training Time	Fastest	Moderate	Slowest (due to self-att
Inference Speed	Fast	Moderate	Slow (sequential decod
Model Complexity	Low	Medium	High
Interpretability	Low (Black-box)	Medium (Attention weights)	High (Attention maps)

THANK YOU