Comparative Study of Sequential and Attention-Based Neural Architectures for Sentiment Analysis

Name: Sanjana S. Joshi
Roll No: MS2405

October 2025

Proposed Work

Sentiment analysis is a central problem in Natural Language Processing (NLP), aimed at automatically determining the emotional tone of textual data [1]. Over the years, several neural network architectures have been proposed. Sequential models such as Recurrent Neural Networks (RNNs) and their variants - Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) - have shown promising results in capturing temporal dependencies [4]. Convolutional Neural Networks (CNNs) have also been applied to sentence classification, efficiently capturing local phrase-level features [3]. However, these models often struggle with long-range dependencies and parallel computation limitations.

The Transformer architecture, based on self-attention mechanisms, addresses these limitations by allowing models to process sequences in parallel and capture global dependencies [5]. While large pre-trained Transformers such as BERT achieve state-of-the-art performance, there is limited research comparing them directly with traditional recurrent models on multiple sentiment datasets, especially in terms of interpretability and efficiency.

Problem Statement

Given a collection of text reviews, the goal is to classify each review as **positive** or **negative** in sentiment.

- **Input:** Raw textual review (e.g., from IMDB dataset)
- **Output:** Sentiment label (1 for positive, 0 for negative)

Formally, given an input sequence of words

$$X = [x_1, x_2, \ldots, x_T],$$

the model should learn a mapping

$$f : X \to y,$$

where

$$y \in \{0, 1\}$$

represents the sentiment class.

Research Gap and Motivation

Although several studies exist on individual architectures, a comprehensive, controlled comparison between recurrent models (RNN, LSTM, GRU) and fully implemented Transformer models (encoder-decoder) is lacking. Most prior works either focus solely on pre-trained Transformers or only sequence-based models without attention mechanisms. Furthermore, interpretability analysis across models and datasets is often missing.
This project addresses these gaps by:

- Implementing RNN, GRU, LSTM, and a full Transformer (encoder-decoder) for sentiment classification.
- Evaluating performance across multiple datasets: IMDB Movie Reviews, SST-2, and Amazon Product Reviews.
- Comparing results with pre-trained Transformers like DistilBERT to analyze trade-offs between performance, computational cost, and interpretability.
- Visualizing model behavior using attention heatmaps and saliency techniques to interpret word-level contributions.

Novelty and Proposed Approach

The novelty of this work lies in its unified experimental framework evaluating sequential and attention-based models consistently. Key contributions include:

- Full Transformer implementation from scratch (encoder + decoder) for sentiment classification.
- Multi-dataset evaluation to assess generalization.
- Comparative analysis with pre-trained models (e.g., DistilBERT).
- Interpretability analysis using attention maps and saliency visualizations to explain predictions.

Methodology Overview

The project follows a structured pipeline:

1. **Dataset Collection and Preparation:** IMDB, SST-2, and Amazon Reviews; preprocessing includes lowercasing, tokenization, and padding/truncation.
2. **Embedding Representation:** Use GloVe embeddings or learn embeddings during training.
3. **Model Implementation:** RNN, GRU, LSTM, and full encoder-decoder Transformer using PyTorch.
4. **Training and Optimization:** Cross-entropy loss, Adam optimizer, dropout regularization, and early stopping.
5. **Evaluation:** Accuracy, precision, recall, F1-score, confusion matrices, training time, and parameter count.
6. **Interpretability:** Attention heatmaps for Transformer and saliency maps for all models.

Expected Deliverables

- Comparative performance report of RNN, GRU, LSTM, and Transformer across three sentiment datasets.
- Implementation of a full Transformer (encoder-decoder) for sentiment classification.
- Visualization of attention and saliency maps for interpretability.
- Research report documenting empirical results, insights, and recommendations.

References

[1] Sentiment analysis. `https://en.wikipedia.org/wiki/Sentiment_analysis`. Accessed: 2025-10-04.

[2] Nicholas J. Higham. *Handbook of Writing for the Mathematical Sciences.* SIAM, 1998.

[3] Yoon Kim. Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882*, 2014.

[4] Richard Socher, Alex Perelygin, Jean Wu, Jason Chuang, Christopher D. Manning, Andrew Ng, and Christopher Potts. Recursive deep models for semantic compositionality over a sentiment treebank. *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, 2013.

[5] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, 2017.