

Title: A Lightweight Study on Attention Mechanisms in Neural Networks

Abstract:

Attention mechanisms allow neural networks to dynamically focus on relevant parts of input data. This study evaluates scaled dot-product attention in sequence modeling tasks and analyzes its computational efficiency and contextual learning capabilities.

Introduction:

Traditional recurrent models struggle with long-range dependencies. Attention improves representation learning by assigning weights to important tokens in the input sequence.

Methodology:

Implemented scaled dot-product attention.

Compared attention-based model with baseline LSTM.

Evaluated on a small text classification dataset.

Results:

Improved contextual understanding.

Reduced dependency loss in longer sequences.

Faster convergence compared to vanilla LSTM.

Conclusion:

Attention significantly enhances sequence modeling and forms the foundation of transformer architectures.