

Exploring Transformer Architectures for NLP

Jane Doe, John Smith

Proceedings of NeurIPS 2024

Published: 2024

DOI: 10.1234/transformers.2024.001

Abstract

We explore transformer models for natural language processing tasks and analyze their performance.

Intro

Transformers have reshaped NLP by enabling parallel training and long-range dependencies. In this section,

Methodology

We evaluate encoder-only and encoder-decoder variants on translation and question answering. Experiments

Results

Transformer Large outperforms baselines by 5% BLEU on WMT14 and achieves 90% accuracy on SC

Discussion

We discuss limitations including compute costs and data bias.