Advances in Deep Learning for Natural Language Processing

Authors: Smith, J., Johnson, A., Williams, R.

Category: Machine Learning

Abstract

This paper explores recent advances in deep learning architectures for natural language processing tasks. We present a comprehensive survey of transformer-based models and their applications in text classification, machine translation, and question answering systems.

Introduction

Natural language processing has seen remarkable progress with the advent of deep learning. Transformer architectures, introduced by Vaswani et al., have revolutionized the field. These models use self-attention mechanisms to capture long-range dependencies in text. Pre-trained language models like BERT and GPT have achieved state-of-the-art results across numerous NLP benchmarks. Fine-tuning these models on downstream tasks has become the standard approach. Recent work has focused on scaling these models to billions of parameters, leading to emergent capabilities in few-shot learning and reasoning.