Neural Machine Translation

This paper introduces a novel Neural Machine Translation (NMT) approach by jointly learning to align and translate using a soft attention mechanism.

Unlike traditional encoder-decoder models that encode the entire source sentence into a fixed-length vector, the proposed model (RNNsearch) dynamically selects relevant parts of the source sentence at each decoding step. This allows better handling of long sentences and improves translation quality.

The model achieved performance comparable to traditional phrase-based systems and showed intuitive word alignments in qualitative analysis. This research marks a foundational step in the development of attention-based models in natural language processing, influencing future architectures such as the Transformer.

Key Contributions:
- Proposed a **soft attention mechanism** for better translation alignment.
- Overcame the **fixed-length vector limitation** in traditional encoder-decoder models.
- Improved translation of **longer sentences** with better context handling.
- Achieved **state-of-the-art performance** comparable to phrase-based SMT.
- Paved the way for modern attention-based models in NLP.