**Review of 1st One**

**Attention Is All You Need**

This research paper has been written about a new proposed dominant sequence transduction model, based on complex recurrent or convolutional neural networks that include an encoder and a decoder, as simple network architecture. After training for 3.5 days the model has achieved 28.4 BLEU, over 2 BLEU over the existing best result, on the WMT 2014 English-to-German translation task and 4.18 BLEU, a number between zero and one that measures the similarity of the machine-translated text to a set of high quality reference translations, on the WMT 2014 English-to-French translation task. The authors showed that after transforming the model it generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.