



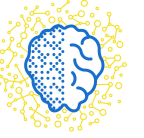
Advances in NLP, 2017

Valentin Malykh,

Neural Systems and Deep Learning lab

Moscow Institute of Physics and Technology

Trends

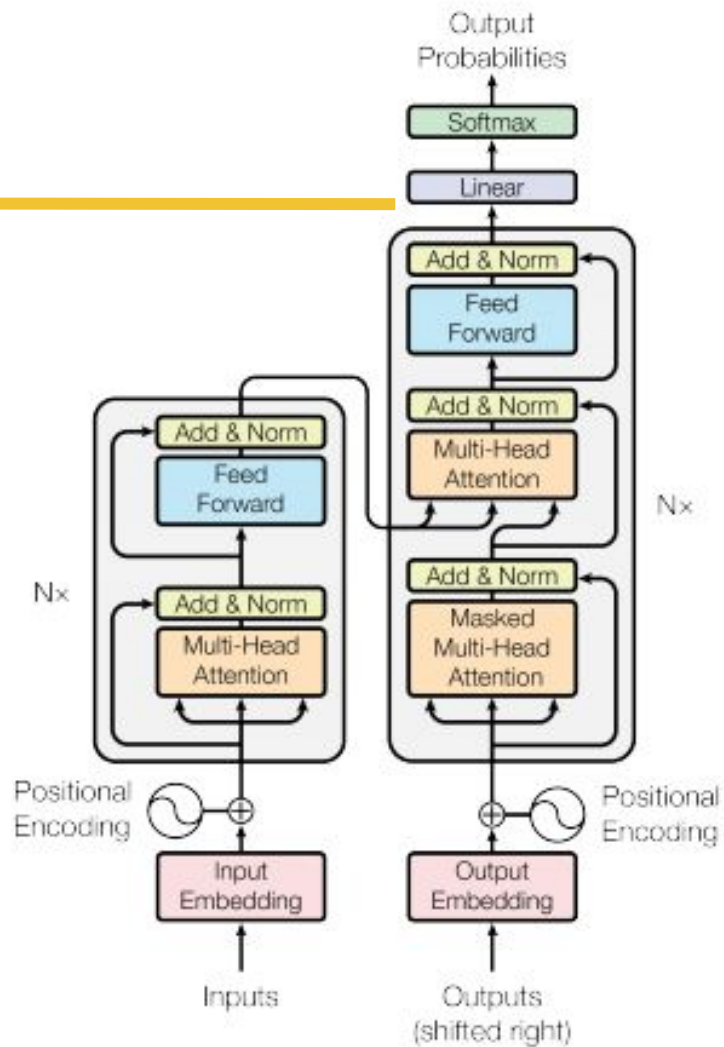


- Faster
- Parallel
- Unsupervised

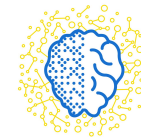
iPavlov.ai

Attention is All You Need

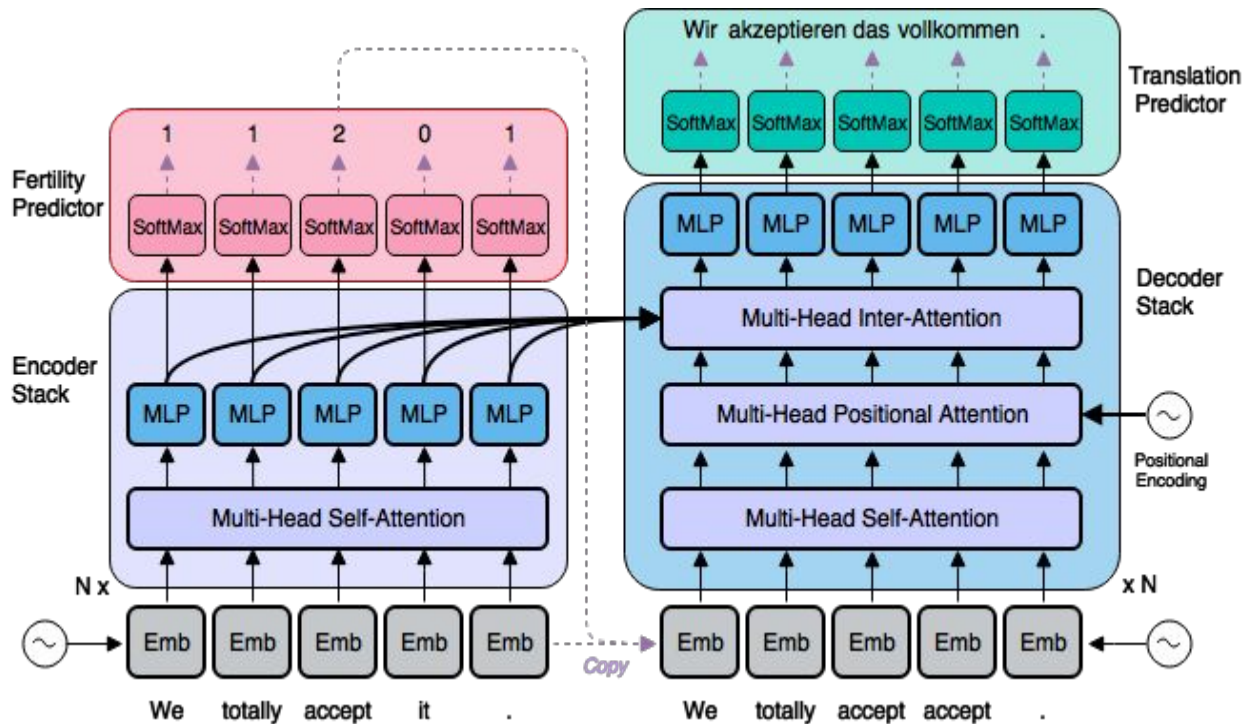
- self-attention
- positional encoding
- masked generation



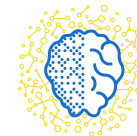
Fully parallel Text-Generation for MNT



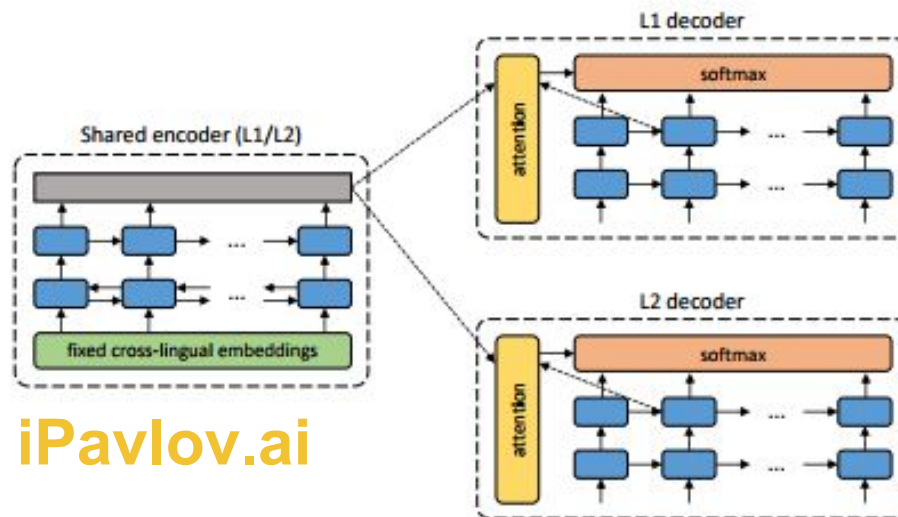
- fertility
- self-attention
- positional encoding



Unsupervised Machine Translation

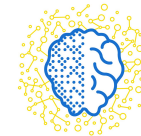


- Unsupervised Neural Machine Translation (arxiv:1710.11041)
- Unsupervised Machine Translation Using Monolingual Corpora Only (arxiv:1711.00043)

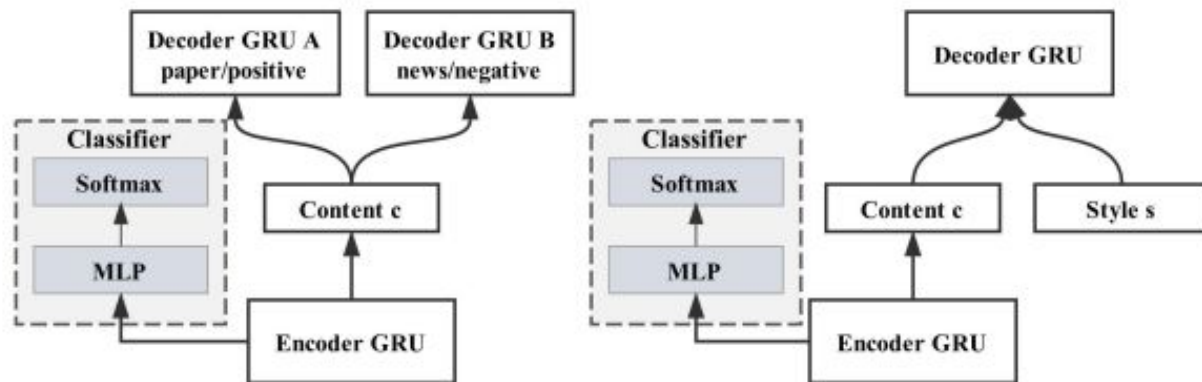


iPavlov.ai

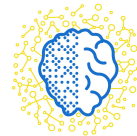
Style Transfer on Texts



- Style Transfer in Text: Exploration and Evaluation (arxiv:1711.06861)
- Toward Controlled Generation of Text (arxiv:1703.00955)



Simple Recurrent Unit



- Training RNNs as Fast as CNNs (arxiv:1709.02755)

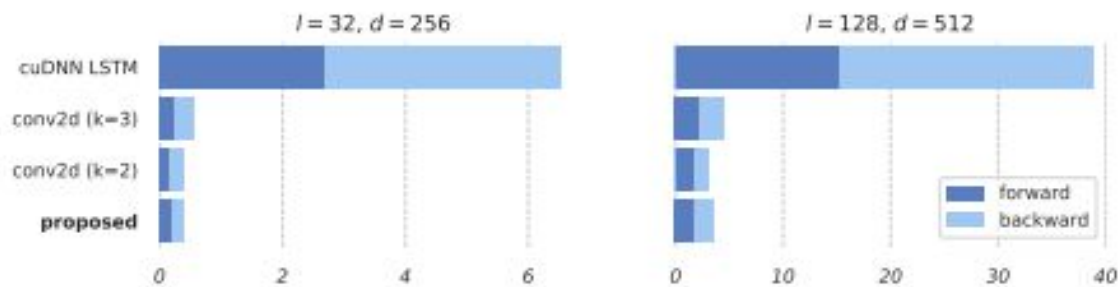
$$\tilde{\mathbf{x}}_t = \mathbf{W}\mathbf{x}_t$$

$$\mathbf{f}_t = \sigma(\mathbf{W}_f\mathbf{x}_t + \mathbf{b}_f)$$

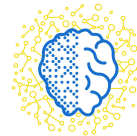
$$\mathbf{r}_t = \sigma(\mathbf{W}_r\mathbf{x}_t + \mathbf{b}_r)$$

$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + (1 - \mathbf{f}_t) \odot \tilde{\mathbf{x}}_t$$

$$\mathbf{h}_t = \mathbf{r}_t \odot \mathbf{g}(\mathbf{c}_t) + (1 - \mathbf{r}_t) \odot \mathbf{x}_t$$



Tools



- AllenNLP
- ParlAI
- OpenNMT

AllenNLP



iPavlov.ai

The background features a network diagram composed of yellow lines connecting various sized white circles. The circles are scattered across the frame, with some having multiple connections, creating a web-like structure. The lines are thin and yellow, while the circles are white with a thin yellow outline.

**Thank you
for your attention!**