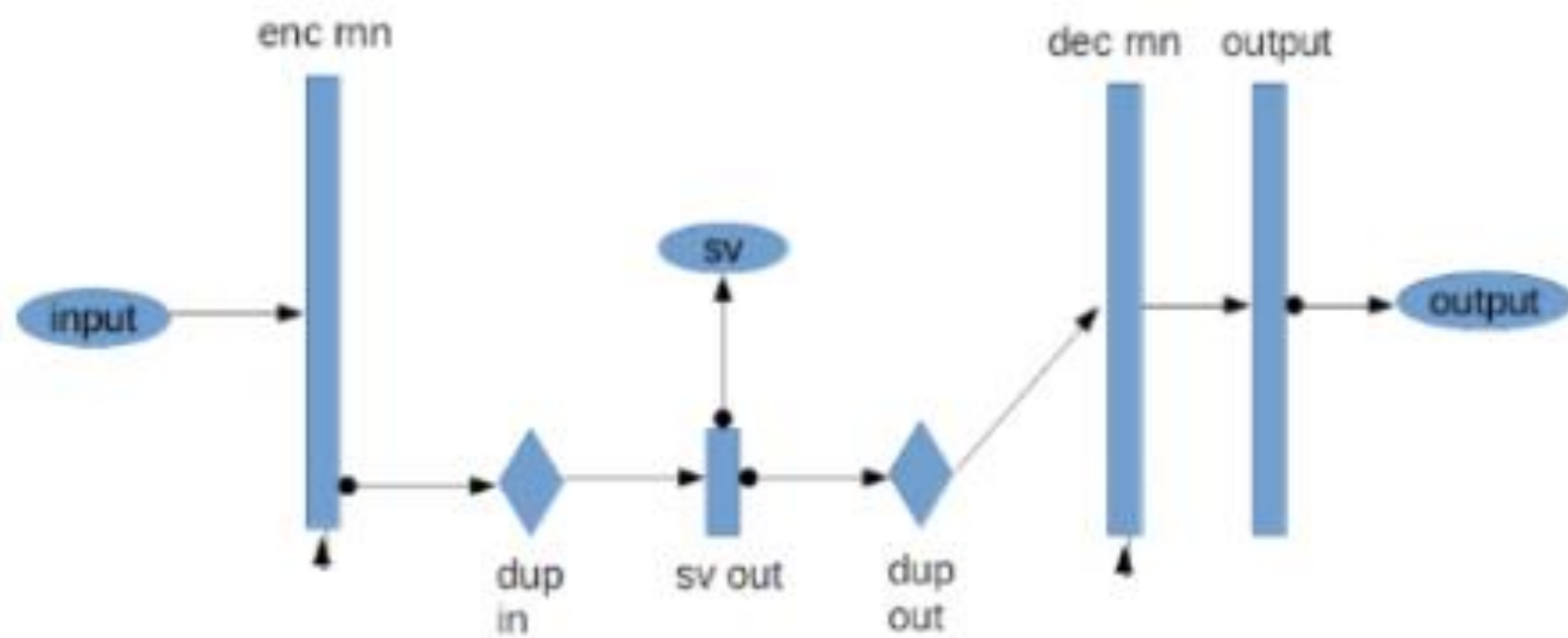


Invertible sentence embeddings

Михеев Артём

RNN



Financial News Headlines Data

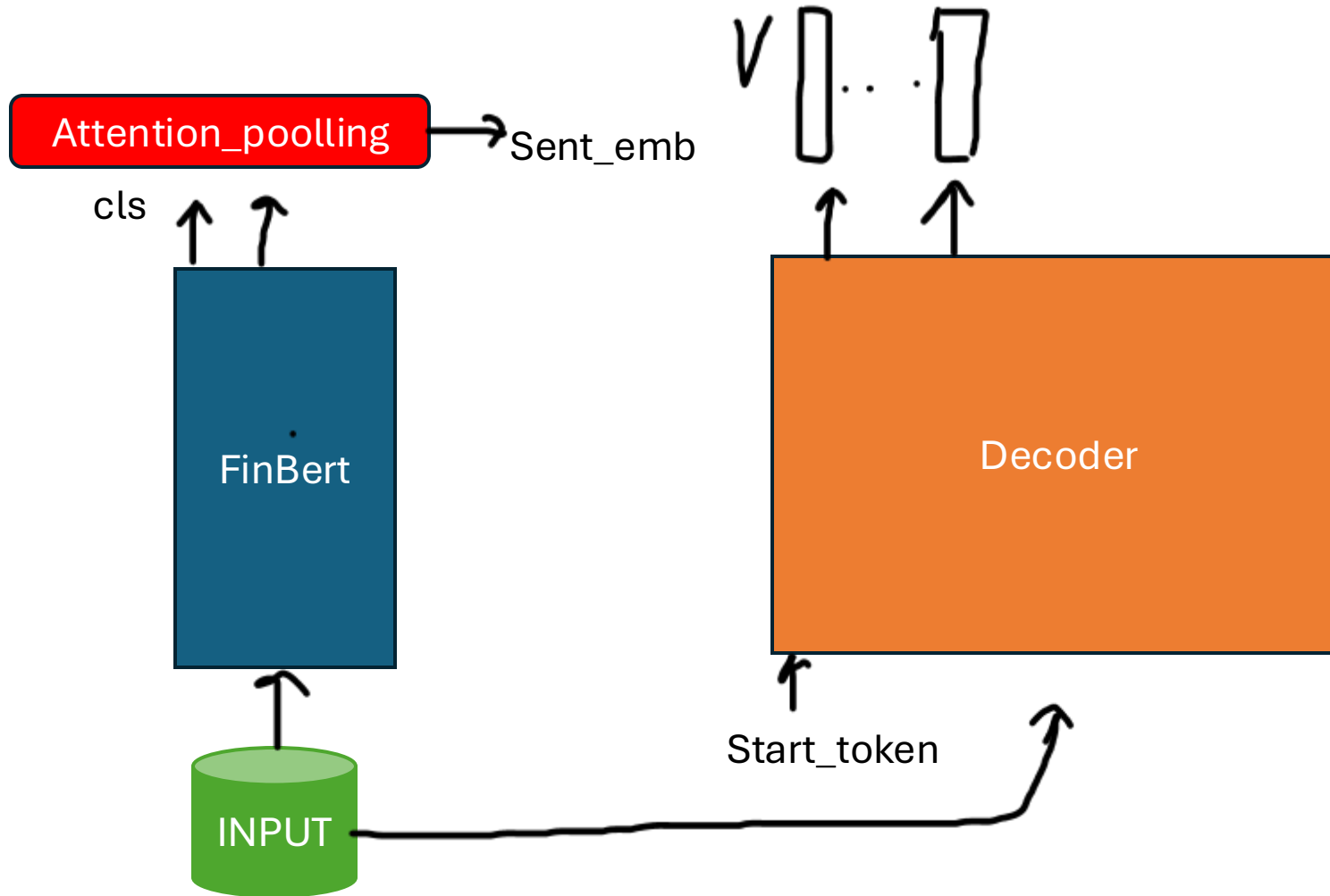
U.S. economy faces significant risks, long road to recovery: IMF staff.

UBS, Morgan Stanley expected to lead Vodafone Tower IPO: sources.

As big U.S. banks let customers delay payments, loan losses remain unclear.

China says it will act to protect its interests after UK Huawei ban.

Attention-архитектуры



AttentionPooling

```
AttentionPooling(  
  (attn): Linear(in_features=768, out_features=1, bias=True)  
  (proj): Sequential(  
    (0): Linear(in_features=768, out_features=2560, bias=True)  
    (1): GELU(approximate='none')  
    (2): LayerNorm((2560,), eps=1e-05, elementwise_affine=True) )
```

```
def forward(self, x, mask=None):  
    attn_scores = self.attn(x).squeeze(-1) # (B, T)  
    if mask is not None:  
        attn_scores = attn_scores.masked_fill(mask == 0, -1e9)  
    attn_weights = torch.softmax(attn_scores, dim=1) # (B, T)  
    return torch.sum(attn_weights.unsqueeze(-1) * self.proj(x), dim=1)
```

Метрики

1)Accuracy

2)Bert-score, "microsoft/deberta-xlarge-mnli"

```
score("The car was damaged in the accident.",  
      "The vehicle got wrecked during the crash.") = 0.91
```

```
Score("The economy is growing rapidly.",  
      'My dog loves running in the park.') = 0.35
```

3)Paraph_sim, 'sentence-transformers/paraphrase-mpnet-base-v2'

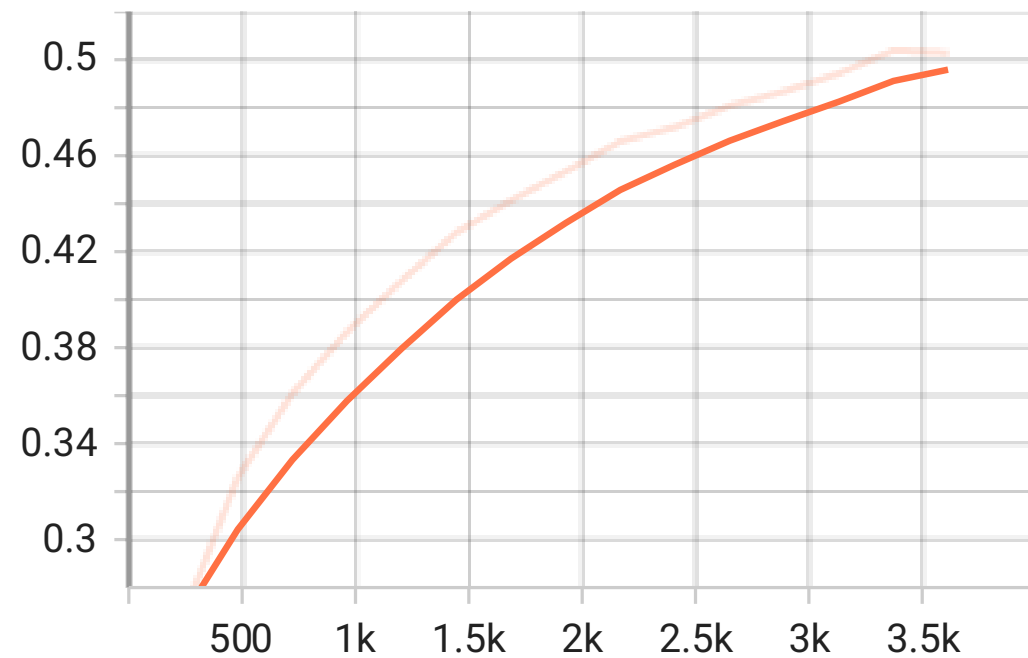
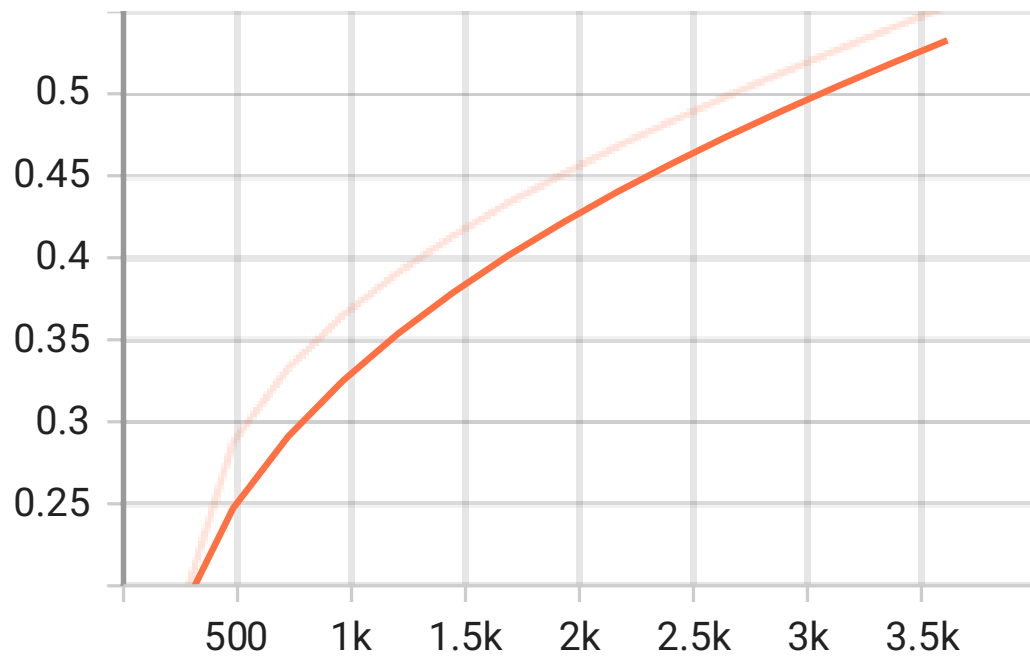
Обучение

dim	Trainable params	Non-trainable params
768	38.5 M	109 M
2560	214 M	109 M

- **Loss:** cross_entropy, sent_sim_loss
- **optimizer:** AdamW
- **lr_scheduler:** ReduceLROnPlateau
- **batch_size:** 128-20

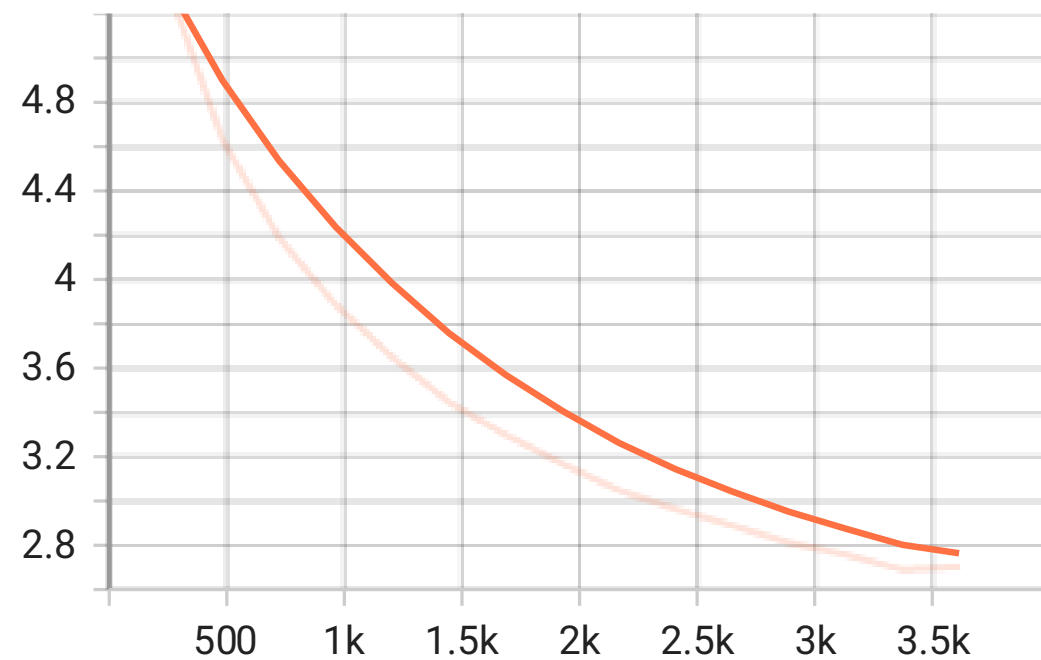
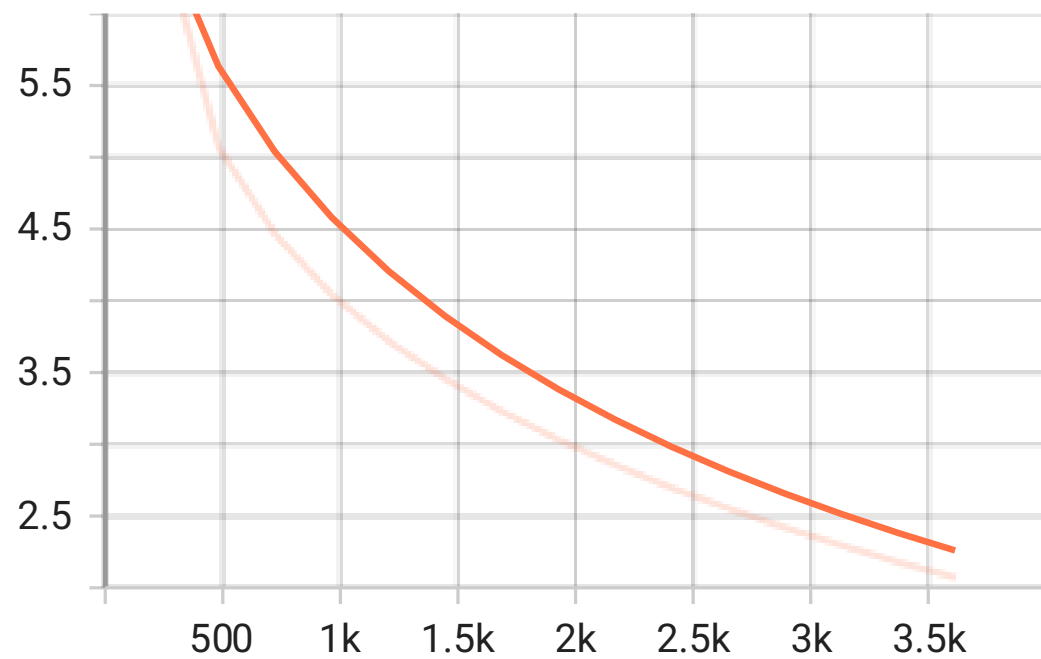
Графики

Accuracy



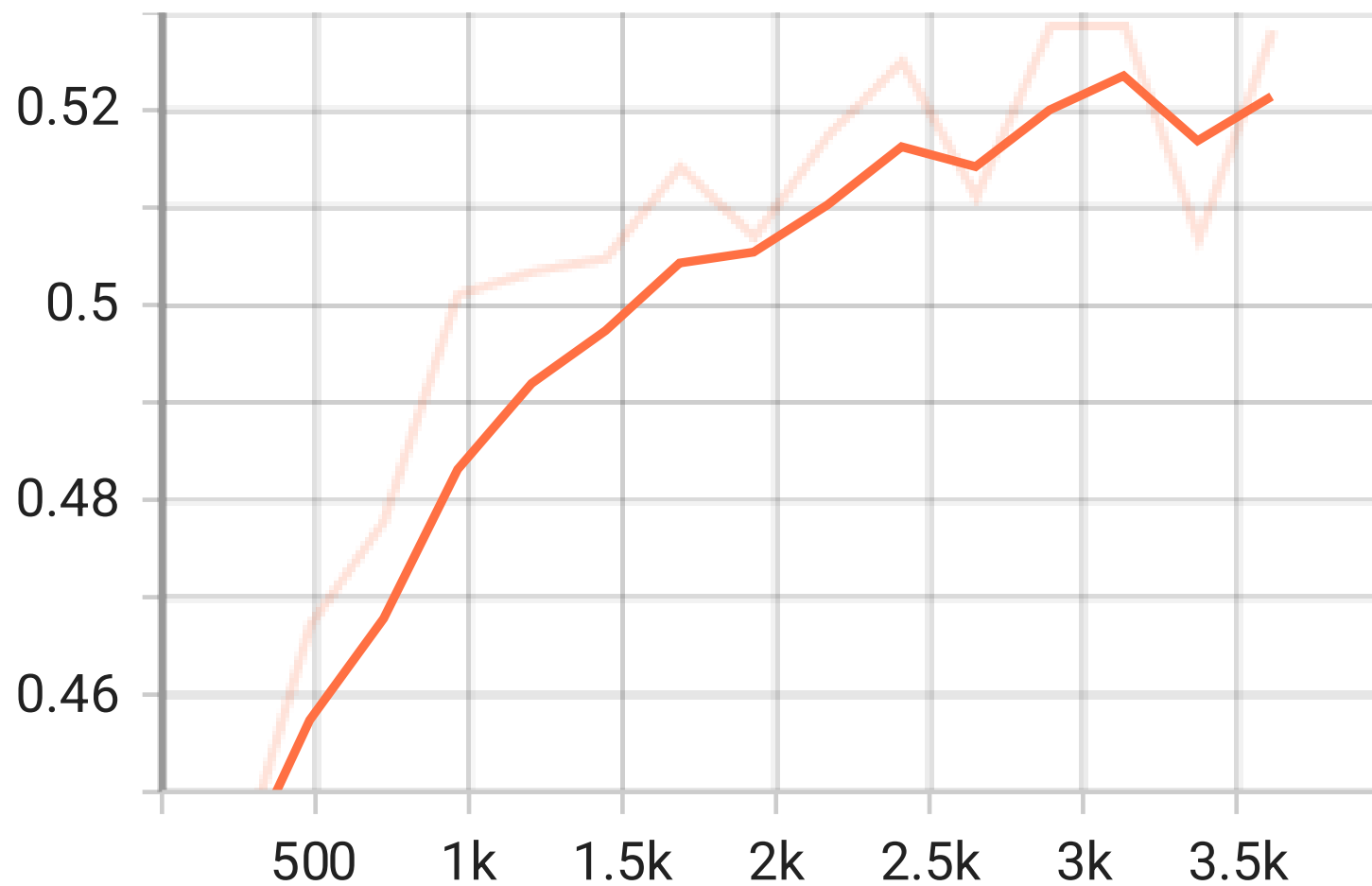
Графики

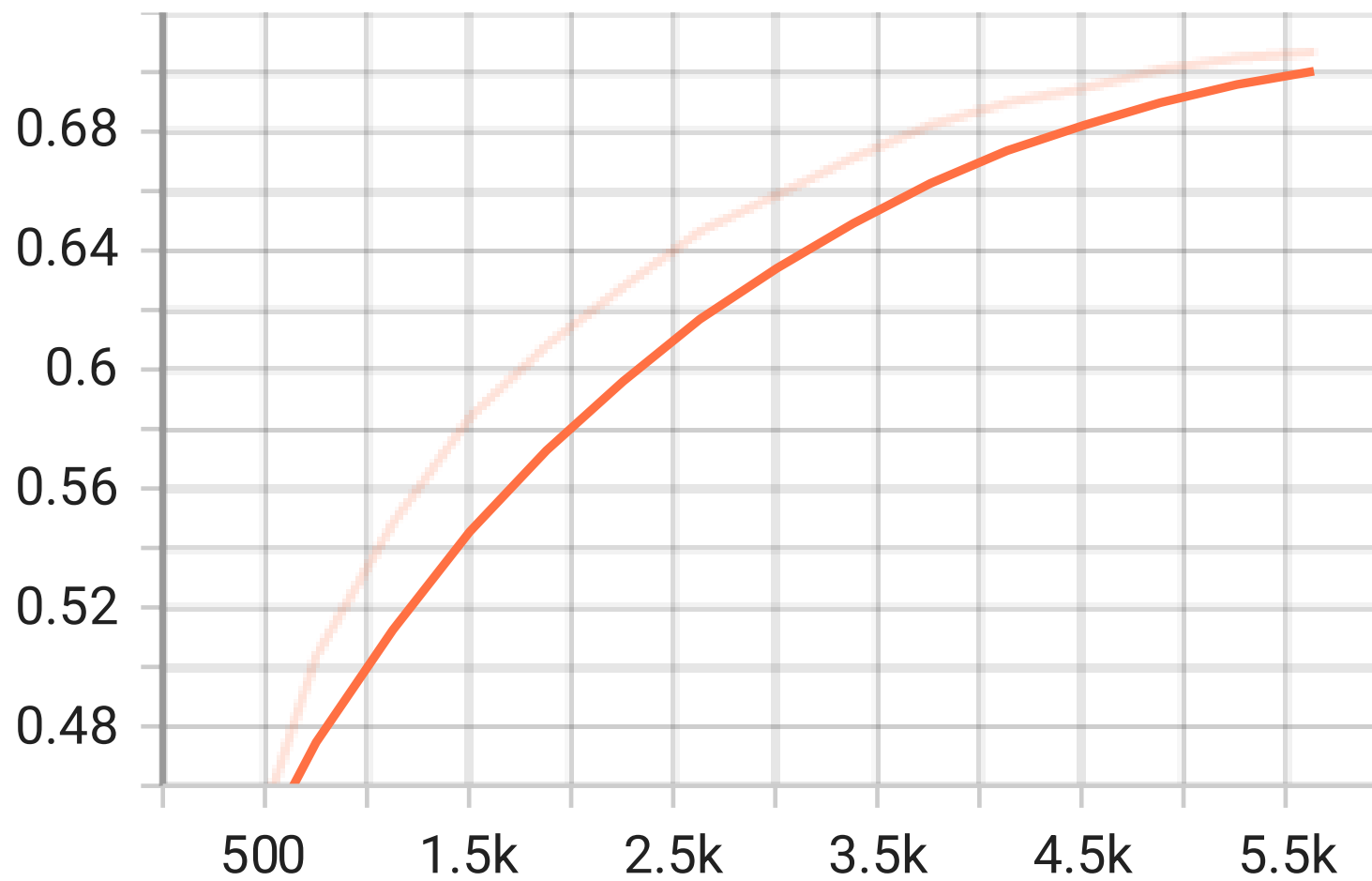
Loss



Графики

Bert-score





Results

model	dim	Valid_accuracy	Bert_score
Encoder_only	1568	0.357	
Cls_decoder	768	0.406	0.58
Decoder_att_pool	768	0.502	0.65
	2560	0.552	0.675

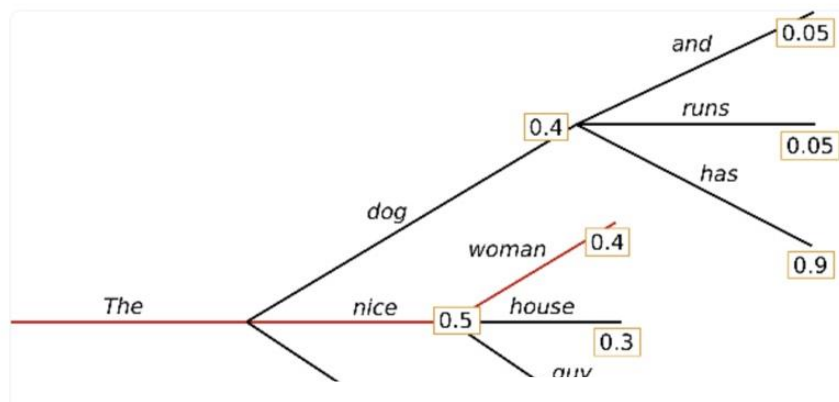
Top-k

Top-k	1	3	5	10	15	25	50	100
Model_768	0.50	0.63	0.68	0.73	0.75	0.78	0.81	0.84
Model-2560	0.55	0.68	0.72	0.77	0.79	0.81	0.84	0.86

Paraph_sim

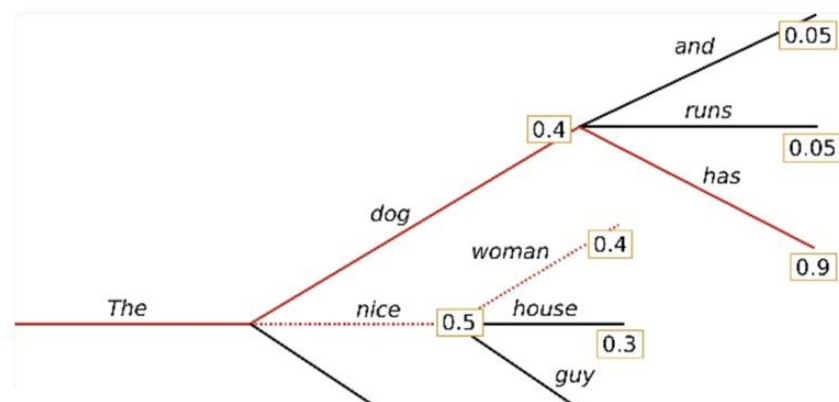
0.64

Greedy



0.71

Beam



Примеры

INPUT	DECODED
consumer giants spurn risks to chase online subscribers	consumer giants spurn consumer risks to online subscribers
t - mobile beats estimates for phone subscribers'	t - mobile beats estimates for phone subscribers
wall street falls as u. s. - china tariffs kick in	wall street falls as u. s. - china tariffs kick in
bayer says monsanto likely kept files on influential people across europe	bayer says monsanto likely put on some countries sick people in europe
wall street rises on economic data, easing trade worries	wall street rises on trade worries, fed data data
uber offers to share journey data with london city planners	uber reveals plans to data with london property

Параллельные статьи

[1\) Cramming 1568 Tokens into a Single Vector and Back Again: Exploring the Limits of Embedding Space Capacity](#)

[2\) Exploring the Latent Capacity of LLMs for One-Step Text Generation](#)