

# **Pluggable Neural Machine Translation Models via Memory-augmented Adapters**

Yuzhuang Xu\*, Shuo Wang\*, Peng Li<sup>™</sup>, Xuebo Liu, Xiaolong Wang, Weidong Liu, Yang Liu<sup>™</sup>

Tsinghua University, Beijing, China / Harbin Institute of Technology, Shenzhen, China

### Introduction



】 Zh: 用墨水书写的谎言一定掩盖个了用鲜皿书写的事实。(w/o style 【 Zh: 墨写的谎言,决掩不住血写的事实。(stylized)



## Method



Our proposed adapters are integrated in transformers as in Fig. A, and the details of adapters are shown in Fig. B. To build the pluggable model, we need 3 steps. First, extract multi-granular phrases use user provided data like Fig. C. Then, continuousize these phrases like Fig. D. Finally, train the adapters using memory drop mechanism like in Fig. E.

#### Experiments

Method		BLEU(↑)			Perplexity(↓)			Classifier Score(↑)		
		En-Zh	Zh-En	Avg.	En-Zh	Zh-En	Avg.	En-Zh	Zh-En	Avg.
Vanilla (Vaswani et al., 2017)		13.7	15.7	14.7	459.6	127.4	293.5	18.0	28.4	23.2
Extreme (Michel and Neubig, 2018)		16.0	17.7	16.9	315.2	113.7	214.5	37.4	43.4	40.4
Adapter (Houlsby et al., 2019)		16.8	19.4	18.1	351.1	121.0	236.1	33.8	58.2	46.0
MT+Rewrite (Syed et al., 2020)		16.3	15.7	16.0	222.4	127.5	349.8	<u>47.0</u>	28.4	37.7
kNN-MT (Khandelwal et al., 2021)		18.9	20.0	19.5	230.7	<u>98.5</u>	164.6	42.2	70.4	56.3
DExperts (Liu et al., 2021)		13.8	15.9	14.9	467.0	127.3	297.2	18.4	31.4	24.9
ChatGPT (OpenAl, 2022)		20.0	13.6	16.8	620.0	131.8	375.9	41.4	24.6	33.0
Memory-augmented Adapter		20.8	21.1	21.0	257.8	110.9	184.3	47.0	69.4	58.2
+ kNN decoding		21.3	21.8	21.6	199.6	95.1	147.4	53.2	85.2	69.2
IT Medical Low Keyen									A.,	
Method	الملاحد		Med	lical	L	aw	KO	ran	AVg	g.
	valid	test	valid	test	valid	test	valid	test	valid	test
Vanilla	28.1	28.4	26.4	27.6	36.2	35.9	10.9	11.5	25.4	25.9
Adapter	30.9	30.5	26.8	28.0	36.0	35.6	12.9	13.5	26.7	26.9
kNN-MT	28.8	29.2	30.0	32.3	38.3	38.4	14.6	15.1	27.9	28.8
Memory-augmented Adapter	31.2	31.1	30.0	32.0	37.5	37.3	14.7	15.3	28.4	28.9
+ kNN decoding	30.5	31.4	31.3	33.5	38.8	38.6	15.7	16.5	29.1	30.0

#### Conclusion

we propose a useful memorybased adapter to build pluggable NMT models, which can let the users customize the generation behavior of NMT models better by simply providing some text samples.



## Code % and data ::

https://www.github.com/xuyuzhuang11/StyleMT Email 🖬 to us:

xyz21thu@gmail.com lipeng@air.tsinghua.edu.cn liuyang2011@tsinghua.edu.cn