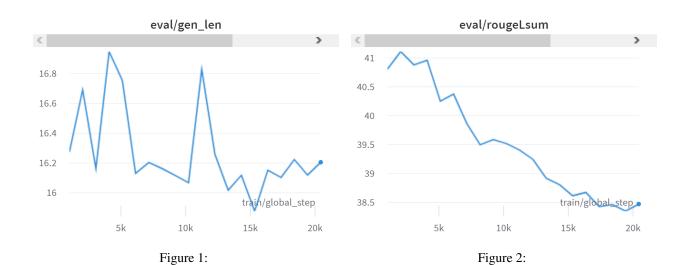
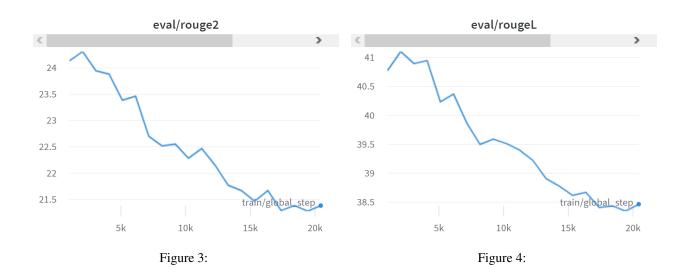
Bart-large 17 Generation

Philipp Leon Alfred Meier meier@cl.uni-heidelberg.de

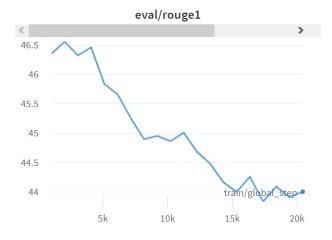
1 Validation & Training





2 References

3 [1] Weights and biases.

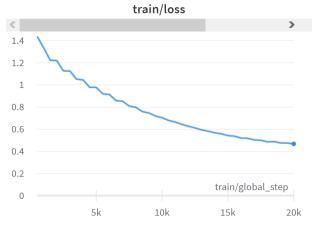


eval/samples_per_second

28
27
26
25
24
23
train/global_step
5k 10k 15k 20k

Figure 5:

Figure 6:



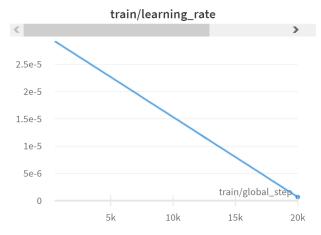
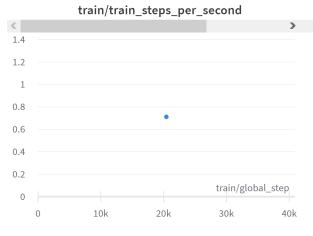


Figure 7:

Figure 8:



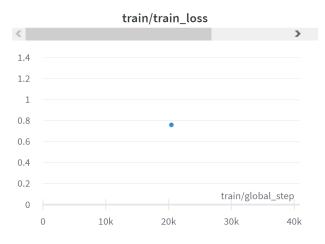


Figure 9: Figure 10:

