

Hyperparameter	Value
seeds	[124, 153, 6582]
num_train_tokens	400M
train_batch_token_num	4096
val_batch_token_num	8184
num_val_batches	30
dict_size	16384
dec_init_norm	0.08
enc_dtype	fp32
lr	5e-05
l1_warmup_pct	0.05
l1_coeff	2
beta1	0.9
beta2	0.999

Table 1: **Hyperparameters for Crosscoder Training.**