Model Depths	Model Dimensions	Attention Heads	Feedforward Dimensions	Avg Train Loss	Avg Eval Loss	BLEU Score:
1	128	2	256	3.970991542	2.268399085	0.4436927704
1	128	2	512	3.906459141	2.169548203	0.4416457332
1	128	4	256	3.93051835	2.20472017	0.4438931079
1	128	4	512	3.900031891	2.150089779	0.4328845596
1	128	8	256	4.02733797	2.287385284	0.4423076042
1	128	8	512	3.952619552	2.204741252	0.4433976611
1	256	2	256	3.11737978	1.489647779	0.4405837328
1	256	2	512	3.090079197	1.48958257	0.444063526
1	256	4	256	3.088508366	1.476128007	0.4421515925
1	256	4	512	3.020979488	1.460565708	0.43263386
1	256	8	256	3.108861851	1.546797428	0.4401196582
1	256	8	512	3.025032981	1.453009332	0.443064364
2	128	2	256	3.941444176	2.102569385	0.4452755567
2	128	2	512	3.912578903	2.071726013	0.4393332381
2	128	4	256	3.960004787	2.151540046	0.4423241232
2	128	4	512	3.911880271	2.097976204	0.442573481
2	128	8	256	4.012285386	2.29237963	0.4452366923
2	128	8	512	3.960776404	2.178096042	0.439568621
2	256	2	256	3.124423238	1.407962739	0.4400710842
2	256	2	512	3.009273465	1.294158987	0.438447103
2	256	4	256	3.070425617	1.367669803	0.4410021554
2	256	4	512	**2.968127032**	1.315076987	0.4408680048
2	256	8	256	3.201760194	1.454110631	0.4442810803
2	256	8	512	3.018960639	1.389965359	0.4619910877
4	128	2	256	4.001598426	2.085616471	0.463946042
4	128	2	512	4.0084471	2.077900372	**0.4699942842**
4	128	4	256	3.979769057	2.096697046	0.4558262987
4	128	4	512	4.038186418	2.135688846	0.4590492427
4	128	8	256	4.122052225	2.317832324	0.456240629
4	128	8	512	4.18972898	2.361531807	0.4592927017
4	256	2	256	3.349210524	1.445724381	0.4568424835
4	256	2	512	3.119592048	1.346217239	0.4623338607
4	256	4	256	3.366589293	1.413882268	0.4582255632
4	256	4	512	3.135615667	**1.278242031**	0.4581841996
4	256	8	256	3.408131894	1.423185509	0.4482460797
4	256	8	512	3.176792338	1.315288784	0.4537801301