

Model	L	Params	Perplexity
Transformer	8	40M	48.8
Transformer	10	46M	48.5
Transformer	10-B	67M	49.2
Transformer	12	52M	48.1
Tree-T	8	44M	46.1
Tree-T	10	51M	45.7
Tree-T	12	58M	45.6