Model	Test PPL	Hardware
Sigmoid-RNN-2048	68.3	1 CPU
Interpolated KN 5-Gram	67.6	100 CPUs
Sparse Non-Negative Matrix LM	52.9	-
RNN-1024 + MaxEnt 9 Gram Features	51.3	24 GPUs
LSTM-2048-512	43.7	32 GPUs
2-layer LSTM-8192-1024	30.6	32 GPUs
BIG GLSTM-G4	23.3[1]	8 GPUs
LSTM-2048	43.9	1 GPU
2-layer LSTM-2048	39.8	1 GPU
GCNN-13	38.1	1 GPU
GCNN-14 Bottleneck	31.9	8 GPUs