	ppl
Full softmax	144
Importance sampling	166
HSM (frequency binning)	166
D-softmax	195
Adaptive softmax	147
Our full softmax	144.17
Self-organized HSM	144.77