Towards Binary-Valued Gates for Robust LSTM Training
Zhuohan Li, Di He, Fei Tian, Wei Chen, Tao Qin, Liwei Wang, Tie-Yan Liu
Peking University & Microsoft Research Asia

1. Long Short-Term Memory (LSTM) RNN

2. Histograms of Gate Distributions in LSTM

3. Training LSTM Gates Towards Binary Values

4. Gumbel-Softmax Estimator

5. Gumbel-Gate LSTM (G-G LSTM)

6. Experimental Results

7. Histograms of Gate Distributions in $G^2$-LSTM

8. Visualization of Average Gate Values