Supplementary figures for article
The Importance of the Current Input in Sequence Modeling

November 21, 2021

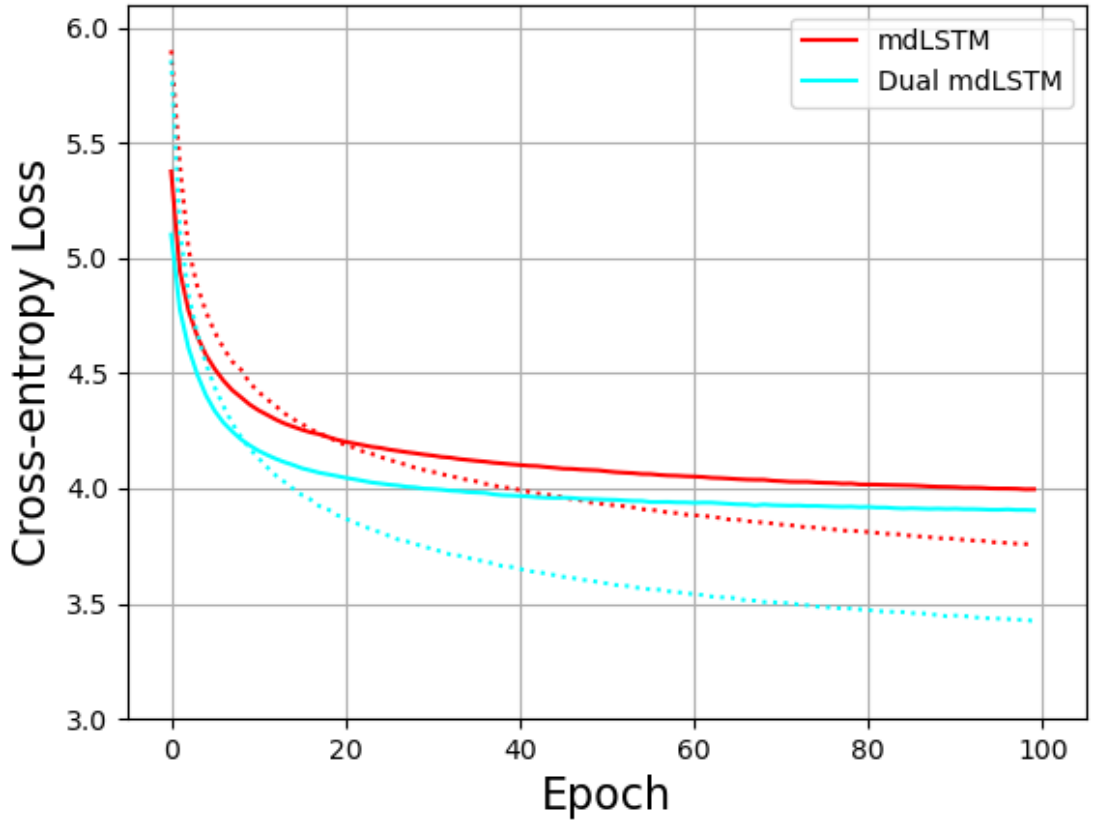

Figure 1: Cross-entropy loss versus epoch for the *mdLSTM* model (baseline, red) and the *Dual mdLSTM* model (cyan) on the PTB dataset. Solid lines are for validation data, dotted lines are for training data.

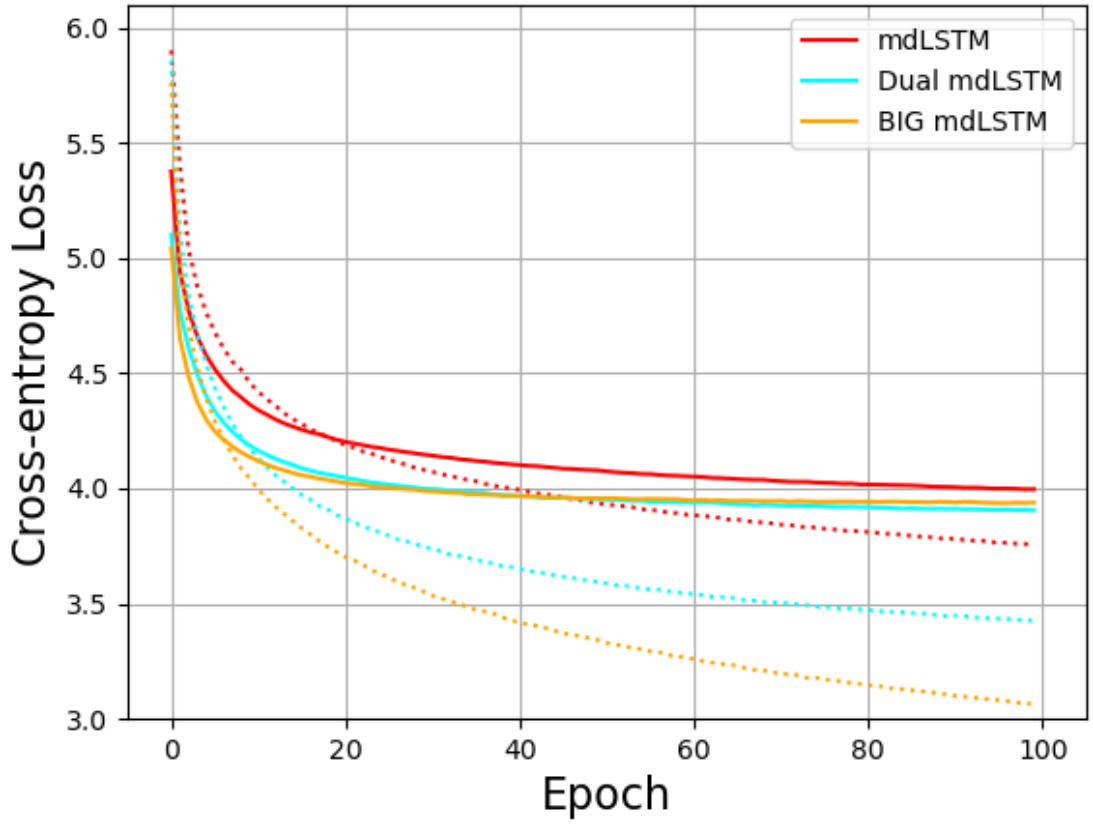

Figure 2: Cross-entropy loss versus epoch for the *mdLSTM* model (baseline, red), the *BIG mdLSTM* model (extended baseline, orange) and the *Dual mdLSTM* model (cyan) on the PTB dataset. Solid lines are for validation data, dotted lines are for training data.

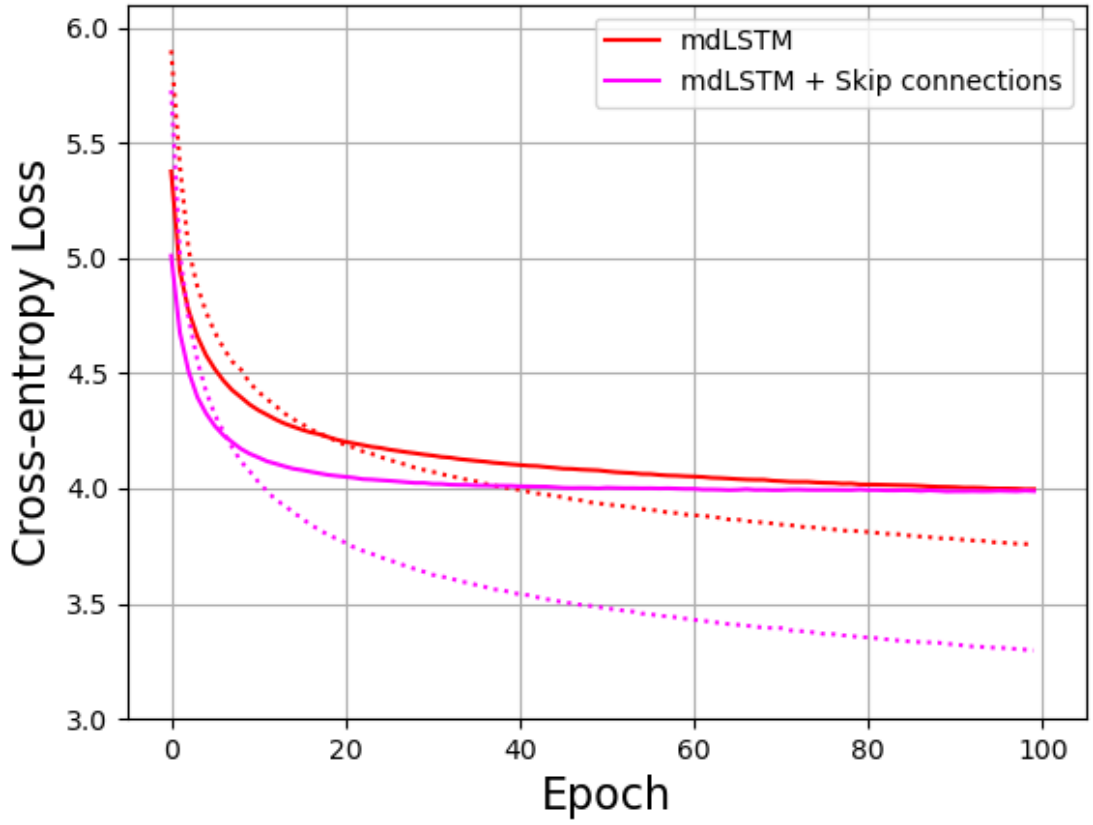

Figure 3: Cross-entropy loss versus epoch for the *mdLSTM* model (baseline, red) and the same model with skip connections (magenta) on the PTB dataset. Solid lines are for validation data, dotted lines are for training data.

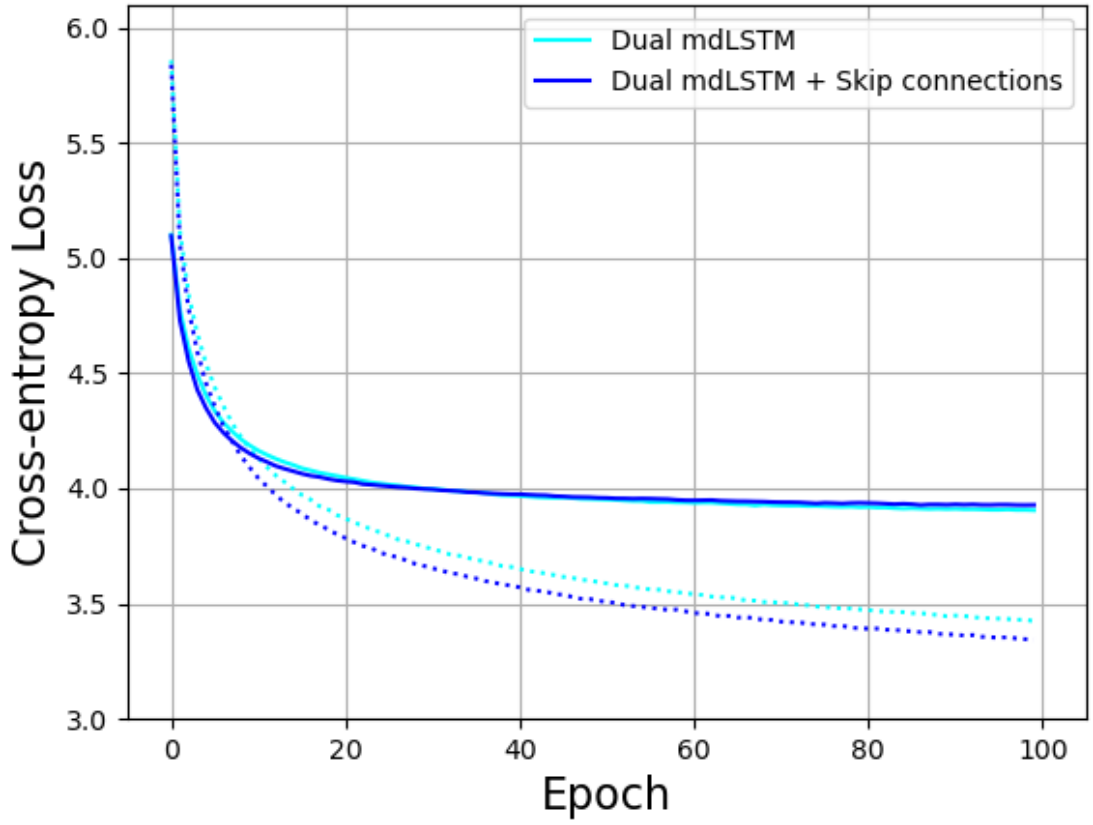

Figure 4: Cross-entropy loss versus epoch for the *Dual mdLSTM* model (cyan) ant the same model with skip connections (blue) on the PTB dataset. Solid lines are for validation data, dotted lines are for training data.