

Arch2 - hidden_size:30; lstm_layer: 8
Avg training and validation loss for each epoch

1e-6+2.5e-1

