

Arch2 - hidden size 30; lstm layer 6  
Avg training and validation loss for each epoch

$1e-5+2.5e-1$

