

Arch2 - hidden size 70; lstm layer 8
Avg training and validation loss for each epoch

$1e-6+2.5e-1$

