

Arch2 - hidden size 50; lstm layer 4
Avg training and validation loss for each epoch

