

Arch2 - hidden size 90; lstm layer 8
Avg training and validation loss for each epoch

