

Arch2 - hidden size 50; lstm layer 8  
Avg training and validation loss for each epoch

