Arch2 - hidden_size 50; lstm_layer 6
Avg training and validation loss for each epoch