Arch2 - hidden_size 70; lstm_layer 6
Avg training and validation loss for each epoch