

Arch2 - hidden_size:10; lstm_layer:8
Avg training and validation loss for each epoch

