Language Model(774M) loss curve