## Training negative log-likelihood

