Marginal log likelihood wrt iterations train_mll valid_mll test_mll -7.55-7.60-7.65-7.700 100 200 300 iter