Marginal log likelihood wrt iterations -13.25 train_mll valid_mll test_mll -13.30-13.35-13.40-13.45100 200 300 iter