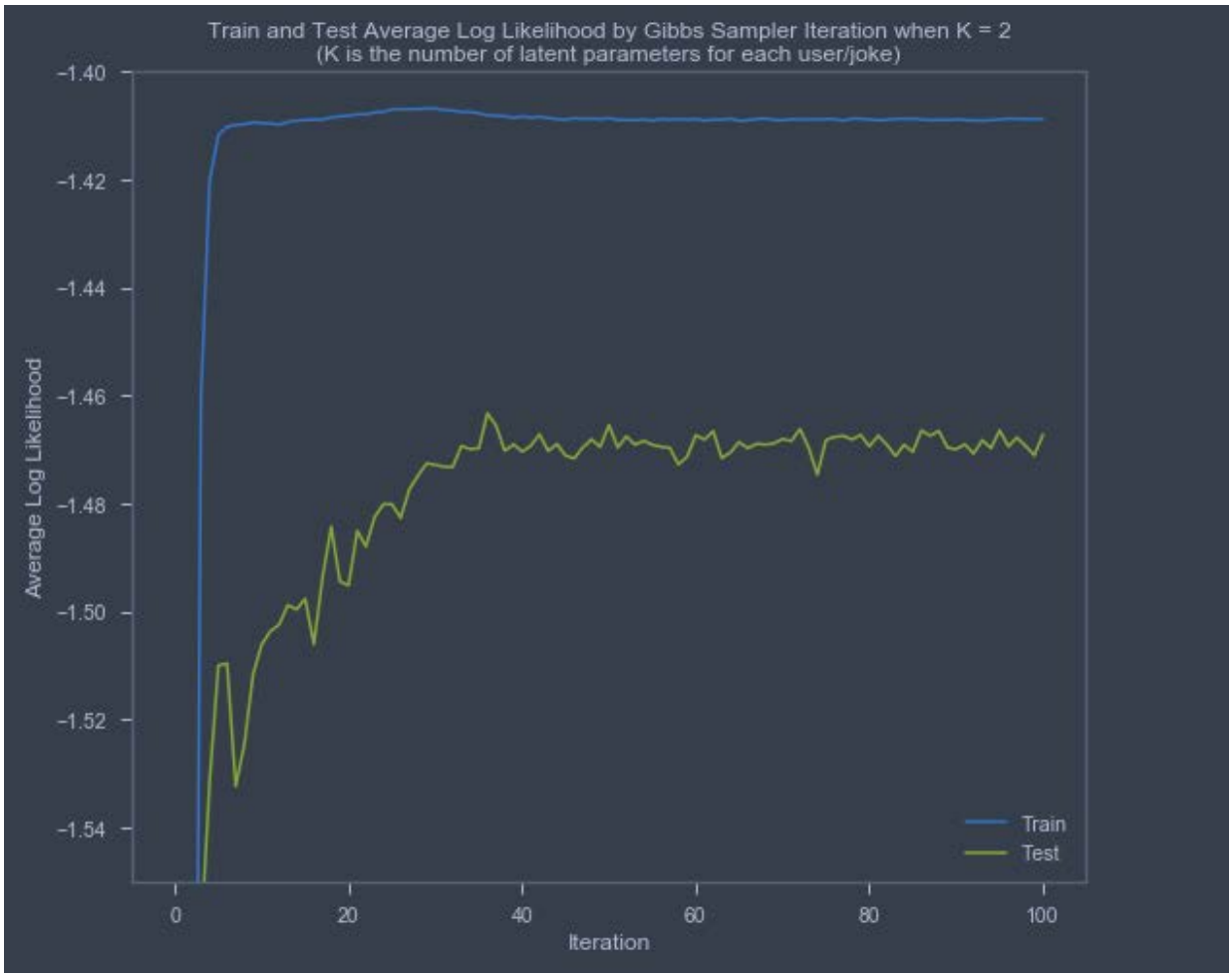
3.3. Note that I plotted average log likelihoods instead of log likelihoods.



3.4. Note that I plotted average log likelihoods instead of log likelihoods.

When doing maximum likelihood estimation in HW3, we saw small improvements in the performance of the model on the training and testing sets as we increased K. Here, when taking a Bayesian approach and using Gibbs sampling, we see small improvements in performance on the training set as we increase K. For the testing set, we see small improvements in performance up through K=4, but then we see small drop-offs in performance as we increase K. These observations are at odds with what we would expect to see when comparing a Bayesian approach to a maximum likelihood approach. However, if we were to increase K further, I think we would be far more concerned about a high degree of overfitting when using the maximum likelihood approach than when using the Bayesian approach with Gibbs sampling.