Programming Assignment 1

Shubham Choudhary, ME15B139
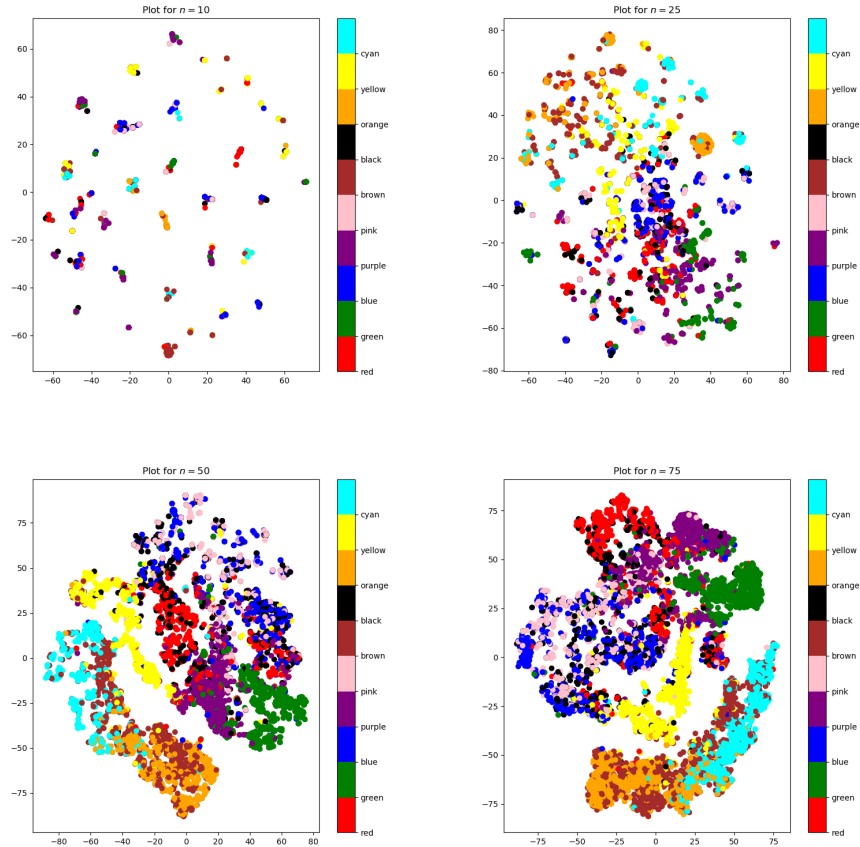Dhruv Chopra, EE16B107
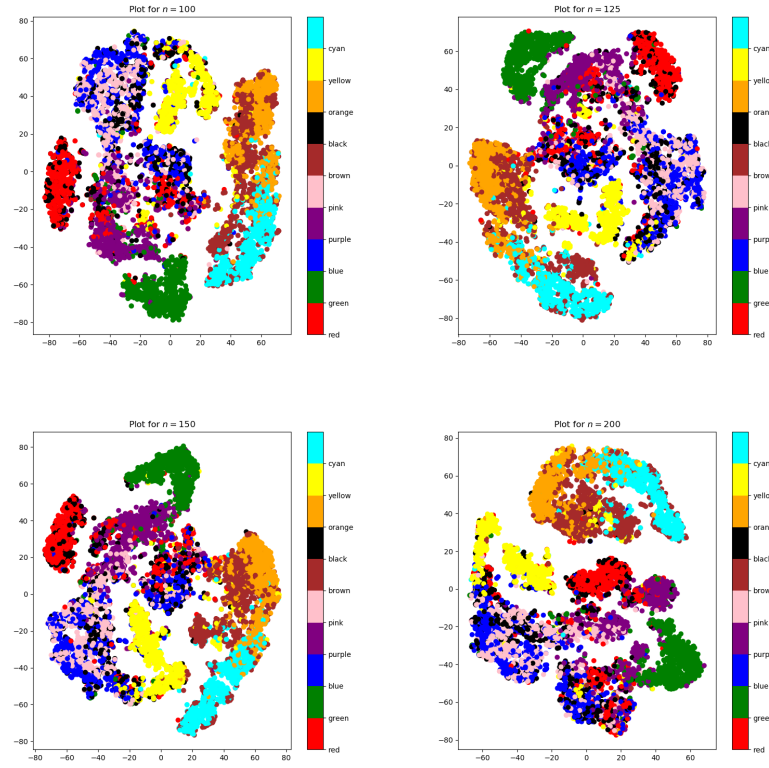
April 26, 2019

1. **Use t-SNE to plot the learned representations in a 2-dimensional space**
   We tried to find the best n for getting the best cluster representation.
   We got best results at $n = 200$. Thus, we infer that more the number of hidden variables, better are the results. However, $n = 100$ is sufficient to get a decent visualization.

2. **Studying the effect of using different values of k**
   We plotted the moving average loss with moving window size to be 1000. This is necessary as there is too much variance in the graph otherwise. As can be observed $k = 20$ converges much faster and gives the lowest reconstruction error for the test data. This can also be inferred from the $T - SNE$ plots for different values of k. We get slightly better clusters for $k = 20$. This is because upon increasing k we run the Gibbs chain for more number of times, and hence we move about slightly in the image space around the training images and thus we obtain a better generalization for the test data.
   The graphs for various values of k are shown below.

Plot for $k = 5$

Plot for $k = 10$

Plot for $k = 20$

Varying k

Varying k
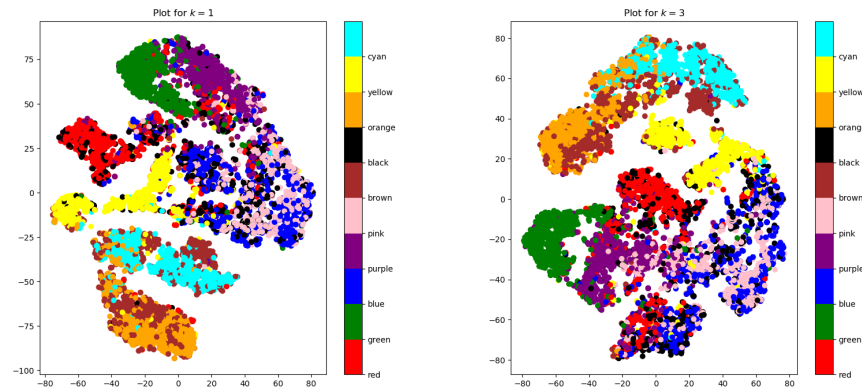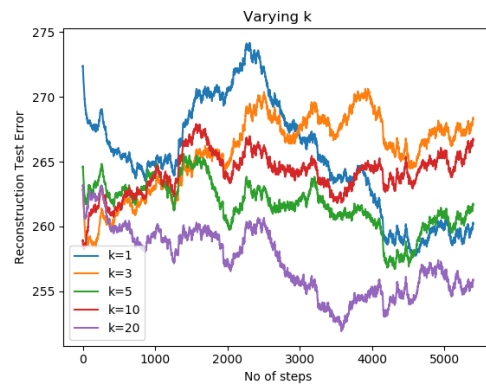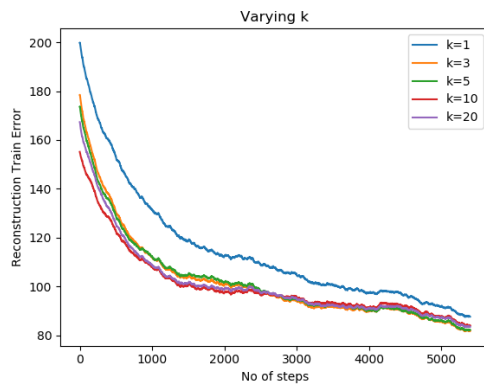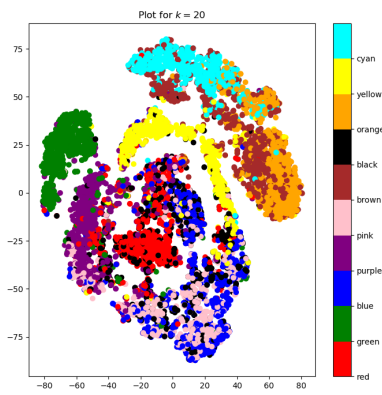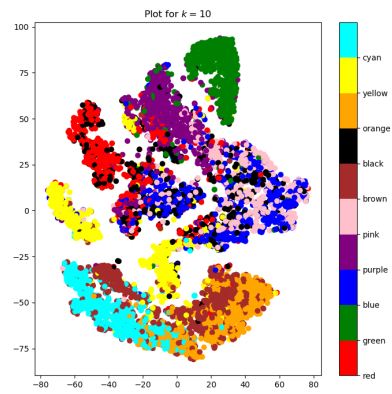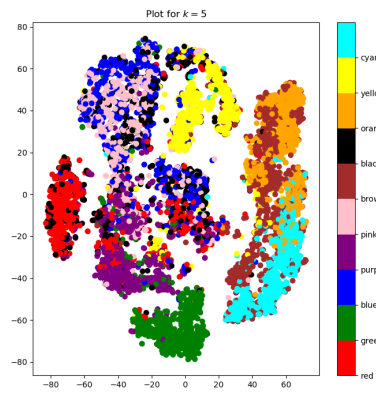
3. **Plot of the samples generated by Gibbs chain**
   We plotted the generated images after each 100 iteration for 6400 iterations. As can be seen, the generated images improve over the iterations.