

The number of clusters should not be too low so that it is not able to cover all the feature points and neither too high so that it becomes overfitting. The variation among all images is also less. This suggests that  $k$  should be small. I took the base value of  $k$  as 10 and computed the accuracy of the model till 45 for different iterations. I was getting the best prediction accuracy near 31 so value of  $k$  is taken to 31 and iterations to be 20.