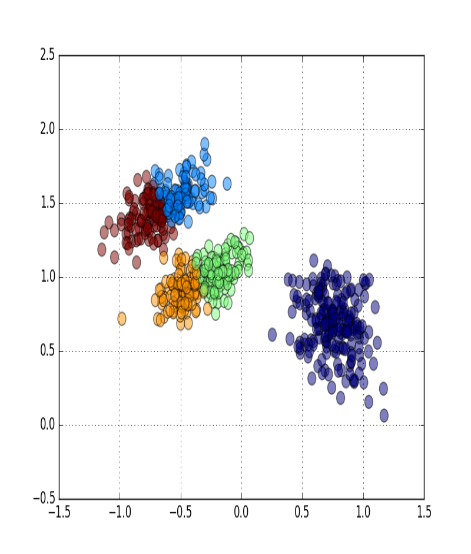
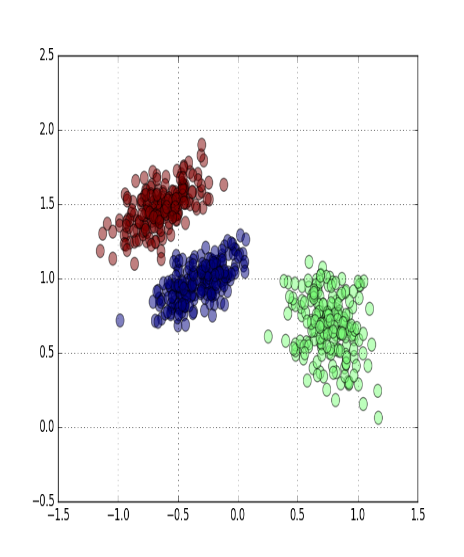
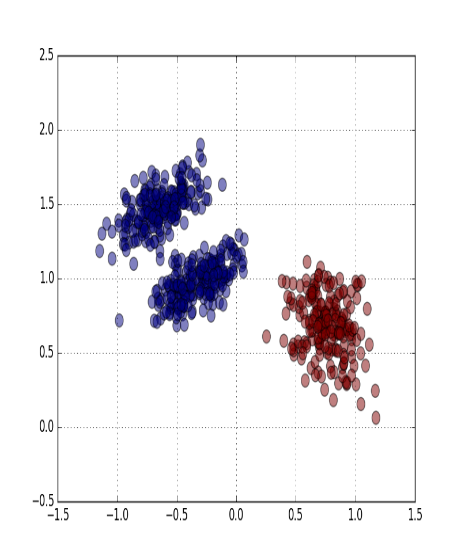
**4. Programming**

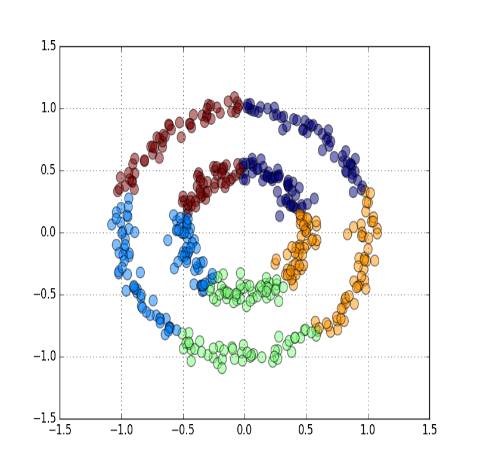
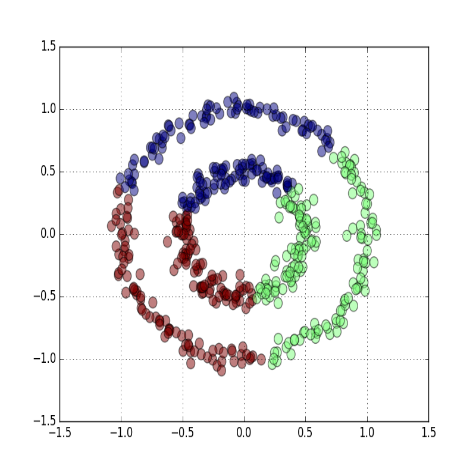
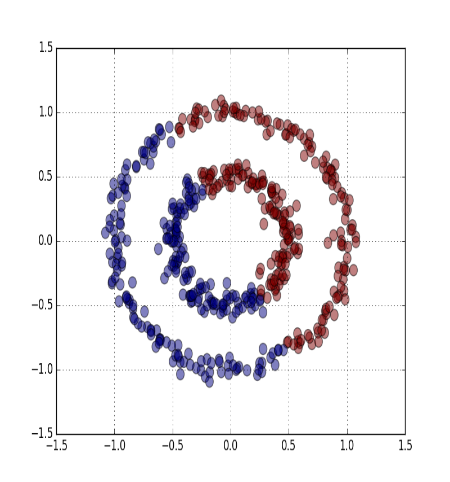
4.2 (a) Implemented k means till no change observed in the clusters assigned.

4.2.a.1

Blob plots for K=2, K=3 and K=5



Circle plots for K=2, K=3 and K=5



4.2. (b) The two circle as shown above are not linearly separable in the original space, and that’s why it is divide into 2 half circles. K- means work on the linear separation of the data points. However, we can transform this into higher dimensional feature space where they might be separable and compute k-means in new feature space.

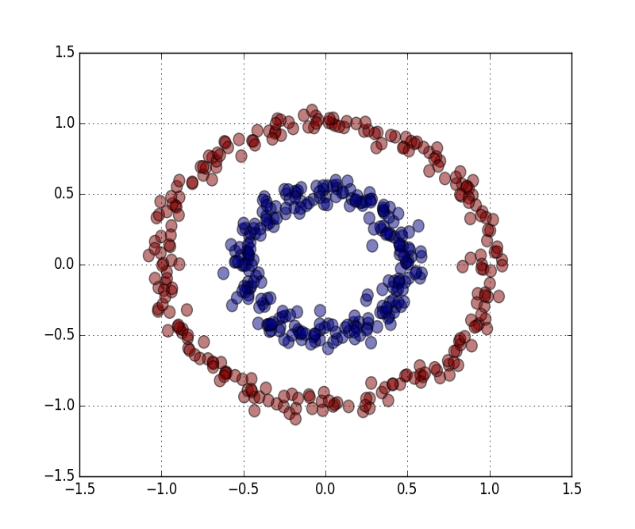
4.3 (a) Experimented with various kernel, as it takes time to converge.

RBF :- = where = 50

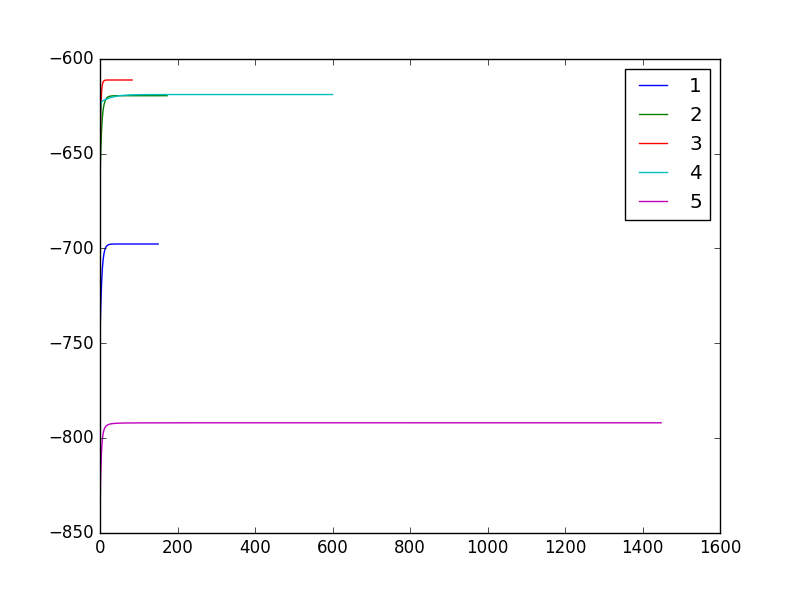
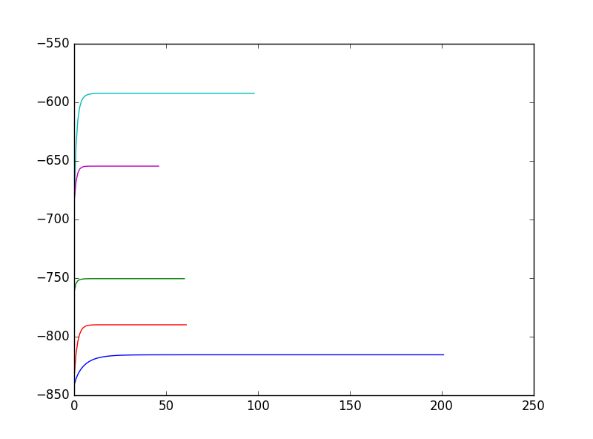
Polynomial :- where c=1 and d=4

For other combination the output was observed to get stuck in the local minimum.

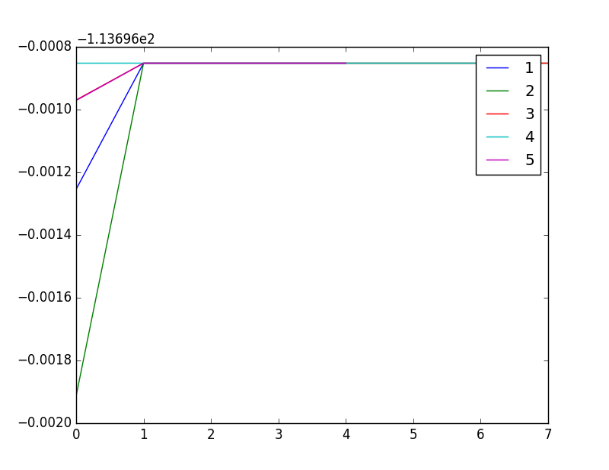
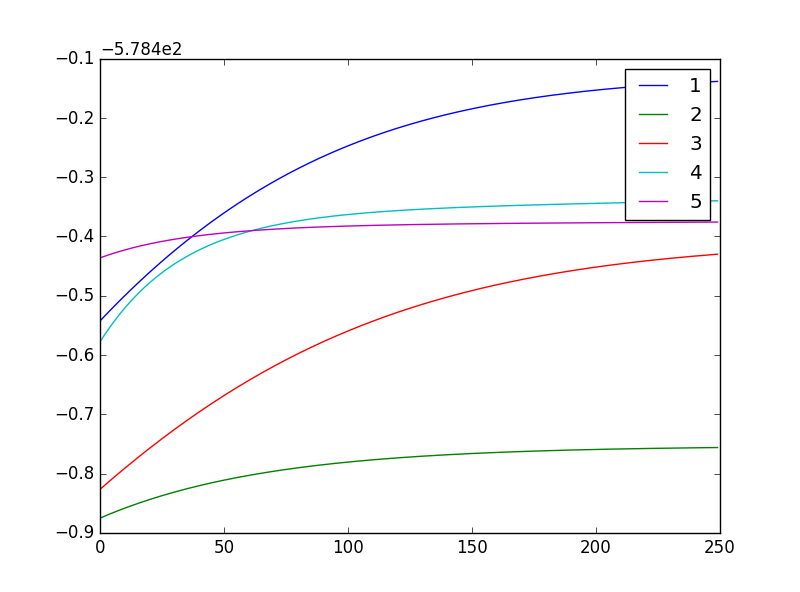
4.3. (b) Following plot was observed for the kernel k means with polynomial kernel , for k=2, c=1 and d=4



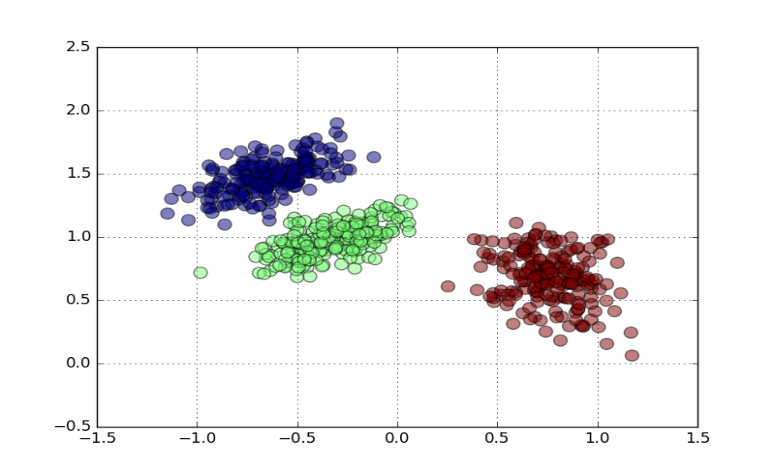
4.4 (a) When randomly initialized the clusters different graph were observed, and takes long time to converge.



H owever, when initialized with output of k-mean , it converges very fast, as shown below.



4.4 (b) Best plot cluster assignments



Best Mean and covariance for the best log likelihood as shown below:

Mean 1= ([-0.63945121, 1.4745009 ]), Covariance 1= [[ 0.03595823, 0.01548446],

[ 0.01548446, 0.01938158]]

Mean 2= ([ 0.75895991, 0.6797701 ]), Covariance 2= [[ 0.02717078, -0.0084006 ],

[-0.0084006 , 0.04044207]]

Mean 3= ([-0.32583659, 0.97128509])], Covariance 3= [[ 0.03603558, 0.01465724],

[ 0.01465724, 0.0162877 ]]