Question 5-

a) To train the network we used 6000 rows and linear kernel, after that the predicted values are:

- The Linear kernel is the simplest kernel function. It is given by the inner product <x,y> plus an optional constant c

$$k(x, y) = x^T y + c$$

Training error (%)	Test error (%)	Num_of_support_vectors
0	2.4	[542,542]

- Parameters used-
  Random state=0
  kernel='linear'

b) Comparing RBF kernel and Polynomial kernel at parameter
   Gamma = 0.001

- Large gamma value means small variance is considered i.e, the points are similar if and only if their variance is small and if gamma value is small i.e high variance is considered and points are considered similar if they are somewhat far from each other

Kernel	Training error (%)	Test error (%)	Num_of_support_vector
RBF	0	15.2	[3000,3000]
Polynomial	0	1.9	[817,938]

- In RBF kernel I used preprocessing of data using standard scalar function of scikit learn, because before preprocessing test error is 50.6 %, So to reduce the test error I used preprocessing of data
- After preprocessing test error in rbf become 15.6%