Report on Assignment 3.2

For SVM classifier:

KERNEL	PENALTY	TEST ERROR RATE
Radial basis function(rbf)	0.1	**73.96**
Radial basis function(rbf)	0.01	**16.50**
Radial basis function(rbf)	1	**13.79**
Linear	0.1	**4.74**
Linear	0.01	**5.57**
Linear	1	**2.46**
Polynomial	0.1	**42.80**
Polynomial	1	**11.39**
Sigmoid	0.1	**88.65**
Sigmoid	1	**32.42**

The best error rate for the above classifier is 2.46 which was getting from parameters such as:

Kernel = linear, and Penalty = 1

For MLP Classifier:

Architecture	Alpha	Activation	TEST ERROR RATE
(100,)	0.0001	Tanh	**2.27**
(100,)	0.0001	Relu	**2.01**
(100,)	0.001	Tanh	**2.23**
(100,)	0.001	Relu	**2.04**
(100,)	0.01	Tanh	**1.87**
(100,)	0.01	Relu	**2.04**
(100,)	0.1	Tanh	**2.10**
(100,)	0.1	Relu	**2.04**
(10,2)	0.0001	Tanh	**9.90**
(10,2)	0.0001	Relu	**10.08**
(10,2)	0.001	Tanh	**9.77**
(10,2)	0.001	Relu	**13.02**
(10,2)	0.01	Tanh	**11.71**
(10,2)	0.01	Relu	**11.13**
(10,2)	0.1	Tanh	**11.08**
(10,2)	0.1	Relu	**11.48**
(5,2)	0.0001	Tanh	**15.47**
(5,2)	0.0001	Relu	**40.99**
(5,2)	0.001	Tanh	**15.98**
(5,2)	0.001	Relu	**16.96**
(5,2)	0.01	Tanh	**18.96**
(5,2)	0.01	Relu	**22.96**

(5,2)	0.1	Tanh	**14.43**
(5,2)	0.1	Relu	**17.17**
(50,30)	0.0001	Tanh	**2.68**
(50,30)	0.0001	Relu	**2.55**
(50,30)	0.001	Tanh	**2.84**
(50,30)	0.001	Relu	**2.79**
(50,30)	0.01	Tanh	**2.42**
(50,30)	0.01	Relu	**2.51**
(50,30)	0.1	Tanh	**2.29**
(50,30)	0.1	Relu	**2.26**

The best error rate for the above classifier is 2.01 which was getting from parameters such as:

Architecture = (100,) , Alpha = 0.0001, and Activation = Relu.

For Gradient Boosting classifier

Number Estimators	Learning rate	Maximum depth	Test Error Rate
50	0.1	1	**17.83**
50	0.1	3	**7.35**
50	0.15	1	**15.30**
50	0.15	3	**6.25**
50	0.2	1	**13.87**
50	0.2	3	**5.67**
70	0.1	1	**15.63**
70	0.1	3	**6.25**
70	0.15	1	**13.59**
70	0.15	3	**5.21**
70	0.2	1	**12.29**
70	0.2	3	**4.99**

The best error rate for the above classifier is 2.01 which was getting from parameters such as:

Number of Estimators = 70, Learning rate = 0.2, and Maximum depth = 3.

K-Means

In this part of assignment, I have compressed the two images and calculated their compression ratios. For Penguin image I have run the algorithm for 14 iteration, and for the Koala image I have run the algorithm for 10 iterations till they converge.

There is a trade off between image quality and degree of compression, as the degree of compression increases the quality of the image gets degrading.

The image quality decreases for lesser values of K since the clusters making the pixel values for the colour vanishing.

For Koala image:

Compression ratios for each iteration

K = 2	K = 5	K = 10	K = 15	K = 20
6.13455	4.61484	4.820599	4.970438	4.964529
5.96263	2.12606	4.719438	4.823249	4.968224
5.99339	4.77826	4.731564	4.957815	5.026043
6.42712	4.1000	4.834148	4.953851	5.001608
6.30556	4.744846	4.838762	4.901331	4.994537
6.07074	4.767706	4.786821	4.807125	4.900962
6.20248	4.346958	4.903886	4.898348	4.892302
5.71844	4.487199	5.081054	4.924204	4.987551
5.91287	4.669791	4.631099	4.833281	5.008923
6.10892	4.492182	4.724607	4.871273	4.998982

Averages

K = 2	K = 5	K = 10	K = 15	K = 20
6.084	4.316	4.805	4.893	4.972

Variance

K = 2	K = 5	K = 10	K = 15	K = 20
0.037	0.571	0.013	0.003	0.001

The good compression is the one whose variance is less, and average is high.

Here for this given image, it is clearly seen that K = 20 will be good value for K because of perfect combination of average and variance as specified earlier.

For Penguin Image:

Compression ratios for each iteration

K = 2	K = 5	K = 10	K = 15	K = 20
4.806613	3.609912	2.635532	2.592524	2.517885
5.113635	2.925501	2.590952	2.525545	2.530656
5.131616	3.858002	2.695771	2.816498	2.575390
5.13375	3.642692	3.037303	2.524315	2.486383
5.091443	3.54719	2.747349	2.45531	2.527933
4.972193	3.712178	2.718344	2.471059	2.5772
5.047599	3.193581	2.668369	2.840275	2.481694
4.936159	3.362607	2.757557	2.716179	2.477505
5.124247	2.856138	2.5635	2.511657	2.703266
5.138838	2.973716	2.622655	2.636827	2.522277

Averages

K = 2	K = 5	K = 10	K = 15	K = 20
5.05	3.013	2.705	2.61	2.541

Variance

K = 2	K = 5	K = 10	K = 15	K = 20
0.011	0.242	0.016	0.018	0.004

The good compression is the one whose variance is less, and average is high.

Here for this given Penguin image, it is clearly seen that K = 2 will be good value for K because of perfect combination of average and variance as specified earlier.