Yahya Alhinai

March 12, 2020

CSCI5521

Problem 1:

A **Model 2:**

$$P(x|) = \frac{1}{2\pi^{\frac{d}{2}} * |(\sum)_i|^{\frac{1}{2}}} \exp(\frac{-1}{2} (x - \mu_i)^T (\sum)_i^{-1} (x - \mu_i))$$

$$L((\sum)_i| x) = \frac{-Nd}{2} log (2\pi) - \frac{N}{2} log |(\sum)_i| - \frac{1}{2}\sum_{t=1}^{N} (x^t - \mu_i)^T (\sum)_i^{-1} (x^t - \mu_i)$$

$$\frac{d}{d(\sum)_i} L((\sum)_i| x) = -\frac{N}{2(\sum)_i} - \frac{1}{2}\sum_{t=1}^{N} (\sum)_i^{-T} (x^t - \mu_i) (x^t - \mu_i)^T (\sum)_i^{-T} = 0$$

$$(\sum)_i = \frac{\sum_{t=1}^{N} (x^t - \mu_i) (x^t - \mu_i)^T}{N}$$

$$S_i = \frac{\sum_{t=1}^{N} (x^t - \mu_i) (x^t - \mu_i)^T}{N}$$

$$S_1 = S_2$$

$$S_1 + S_2 = 2S = \frac{\sum_{t=1}^{N_1} (x^t - \mu_1) (x^t - \mu_1)^T}{N_1} + \frac{\sum_{t=1}^{N_2} (x^t - \mu_2) (x^t - \mu_2)^T}{N_2}$$

$$S = \frac{1}{2}(S_1 + S_2)$$

Model 3:

$$S_i = \alpha_i * I$$

$$L(\alpha_i| x) = \frac{-Nd}{2} log (2\pi) - \frac{N}{2} log |\alpha_i^d| - \frac{1}{2}\sum_{t=1}^{N} (x^t - \mu_i)^T \alpha_i^{-1} I (x^t - \mu_i)$$

$$\frac{d}{d\alpha_i} L(\alpha_i| x) = -\frac{Nd}{2\alpha_i} - \frac{1}{2\alpha_i^2}\sum_{t=1}^{N} (x^t - \mu_i)^T (x^t - \mu_i) = 0$$

$$\alpha_i = \frac{\sum_{t=1}^{N} (x^t - \mu_i)^T (x^t - \mu_i)}{Nd}$$

B

```
------------------------------------------------------------
------------------------------------------------------------
Data Set 1
------------------------------------------------------------
P(C1) = 0.2
P(C2) = 0.8
M1 = 0.6216, 0.5206, -0.5254, -0.1011, -0.0146, 1.0318, 0.2876, -0.0447
M2 = 1.0488  0.2043  0.8059  0.0749  0.5372  0.6031  0.6966  1.0544
-------------------------
Data Set 1 - Model 1
-------------------------
Error Rate = 0.2900
S1 =
8.3481  -4.6489  -4.3645  -3.8639  3.3711  -4.6298  0.1901  5.3910
-4.6489  9.1407  -2.5567  4.2312  -0.7324  4.0982  3.3851  -4.5457
-4.3645  -2.5567  12.8978  3.0713  -2.3547  -1.9411  -5.8484  0.4108
-3.8639  4.2312  3.0713  7.4962  -1.4129  2.8157  -0.0400  -2.8332
3.3711  -0.7324  -2.3547  -1.4129  5.6325  -4.5224  -1.4585  2.3261
-4.6298  4.0982  -1.9411  2.8157  -4.5224  8.5332  4.0668  -6.2549
0.1901  3.3851  -5.8484  -0.0400  -1.4585  4.0668  6.0568  -1.9673
5.3910  -4.5457  0.4108  -2.8332  2.3261  -6.2549  -1.9673  7.0420


S2 =
6.9806  -2.2971  -2.9607  -1.7056  4.5878  -2.2326  -0.3004  2.2206
-2.2971  7.5959  -3.0854  1.0028  1.3683  0.7426  2.4512  -1.8130
-2.9607  -3.0854  9.3598  1.8725  -3.4296  -1.0595  -3.1999  2.5604
-1.7056  1.0028  1.8725  5.2913  0.5197  0.3795  0.0033  -0.0122
4.5878  1.3683  -3.4296  0.5197  7.8898  -2.5563  0.2174  0.9157
-2.2326  0.7426  -1.0595  0.3795  -2.5563  4.2109  1.3974  -2.4591
-0.3004  2.4512  -3.1999  0.0033  0.2174  1.3974  3.5292  -1.7868
2.2206  -1.8130  2.5604  -0.0122  0.9157  -2.4591  -1.7868  4.5089
-------------------------
Data Set 1 - Model 2
-------------------------
Error Rate = 0.2550
S1 =
7.6644  -3.4730  -3.6626  -2.7848  3.9794  -3.4312  -0.0551  3.8058
-3.4730  8.3683  -2.8210  2.6170  0.3179  2.4204  2.9181  -3.1794
-3.6626  -2.8210  11.1288  2.4719  -2.8921  -1.5003  -4.5241  1.4856
-2.7848  2.6170  2.4719  6.3938  -0.4466  1.5976  -0.0183  -1.4227
3.9794  0.3179  -2.8921  -0.4466  6.7611  -3.5393  -0.6206  1.6209
-3.4312  2.4204  -1.5003  1.5976  -3.5393  6.3721  2.7321  -4.3570
-0.0551  2.9181  -4.5241  -0.0183  -0.6206  2.7321  4.7930  -1.8771
3.8058  -3.1794  1.4856  -1.4227  1.6209  -4.3570  -1.8771  5.7755


S2 =
7.6644  -3.4730  -3.6626  -2.7848  3.9794  -3.4312  -0.0551  3.8058
-3.4730  8.3683  -2.8210  2.6170  0.3179  2.4204  2.9181  -3.1794
-3.6626  -2.8210  11.1288  2.4719  -2.8921  -1.5003  -4.5241  1.4856
-2.7848  2.6170  2.4719  6.3938  -0.4466  1.5976  -0.0183  -1.4227
3.9794  0.3179  -2.8921  -0.4466  6.7611  -3.5393  -0.6206  1.6209
-3.4312  2.4204  -1.5003  1.5976  -3.5393  6.3721  2.7321  -4.3570
-0.0551  2.9181  -4.5241  -0.0183  -0.6206  2.7321  4.7930  -1.8771
3.8058  -3.1794  1.4856  -1.4227  1.6209  -4.3570  -1.8771  5.7755
-------------------------
Data Set 1 - Model 3
```

```
-------------------------
Error Rate = 0.6500
σ1 = 7.73624137515677
σ2 = 6.093681509670718


------------------------------------------------------------
------------------------------------------------------------
Data Set 2
------------------------------------------------------------
P(C1) = 0.2
P(C2) = 0.8
M1 = 1.2644, 0.2951, 0.0220, -0.4687, 0.1766, -0.0995, 1.2252, 1.0318
M2 = 1.1875  1.0339  1.2395  0.2855  0.6929  0.0440  0.9373  0.7190
-------------------------
Data Set 2 - Model 1
-------------------------
Error Rate = 0.0450
S1 =
6.6564  0.6897  -4.8294  0.9371  -1.7943  3.1290  2.0019  -4.9913
0.6897  1.7798  1.2671  -0.6797  0.9042  0.6182  -0.2131  -0.9381
-4.8294  1.2671  11.3396  -0.7255  3.9294  2.9642  -1.3406  1.9522
0.9371  -0.6797  -0.7255  6.9097  -2.5162  1.8468  0.7678  -2.0457
-1.7943  0.9042  3.9294  -2.5162  3.2056  -0.1442  -0.2202  2.3499
3.1290  0.6182  2.9642  1.8468  -0.1442  9.0158  1.6407  -6.0387
2.0019  -0.2131  -1.3406  0.7678  -0.2202  1.6407  2.5507  -1.1931
-4.9913  -0.9381  1.9522  -2.0457  2.3499  -6.0387  -1.1931  7.8486


S2 =
8.5976  2.6625  2.2209  -4.0600  -3.6996  -3.6332  -1.9657  -0.9823
2.6625  4.7255  1.2548  -0.2870  -0.0853  -1.5999  0.4875  -1.5974
2.2209  1.2548  3.2014  -1.7007  -0.6558  -0.9575  -0.8843  -0.6915
-4.0600  -0.2870  -1.7007  7.8644  3.9884  1.4676  1.6794  0.6439
-3.6996  -0.0853  -0.6558  3.9884  6.1081  0.8235  2.6251  -0.0603
-3.6332  -1.5999  -0.9575  1.4676  0.8235  4.3934  1.1359  0.7222
-1.9657  0.4875  -0.8843  1.6794  2.6251  1.1359  3.6609  0.4272
-0.9823  -1.5974  -0.6915  0.6439  -0.0603  0.7222  0.4272  2.5246
-------------------------
Data Set 2 - Model 2
-------------------------
       Error Rate = 0.2100
       S1 =
       7.6270  1.6761  -1.3042  -1.5614  -2.7469  -0.2521  0.0181  -2.9868
       1.6761  3.2527  1.2609  -0.4833  0.4095  -0.4909  0.1372  -1.2677
       -1.3042  1.2609  7.2705  -1.2131  1.6368  1.0034  -1.1124  0.6304
       -1.5614  -0.4833  -1.2131  7.3871  0.7361  1.6572  1.2236  -0.7009
       -2.7469  0.4095  1.6368  0.7361  4.6568  0.3397  1.2025  1.1448
       -0.2521  -0.4909  1.0034  1.6572  0.3397  6.7046  1.3883  -2.6583
       0.0181  0.1372  -1.1124  1.2236  1.2025  1.3883  3.1058  -0.3829
       -2.9868  -1.2677  0.6304  -0.7009  1.1448  -2.6583  -0.3829  5.1866


       S2 =
       7.6270  1.6761  -1.3042  -1.5614  -2.7469  -0.2521  0.0181  -2.9868
       1.6761  3.2527  1.2609  -0.4833  0.4095  -0.4909  0.1372  -1.2677
       -1.3042  1.2609  7.2705  -1.2131  1.6368  1.0034  -1.1124  0.6304
       -1.5614  -0.4833  -1.2131  7.3871  0.7361  1.6572  1.2236  -0.7009
       -2.7469  0.4095  1.6368  0.7361  4.6568  0.3397  1.2025  1.1448
       -0.2521  -0.4909  1.0034  1.6572  0.3397  6.7046  1.3883  -2.6583
       0.0181  0.1372  -1.1124  1.2236  1.2025  1.3883  3.1058  -0.3829
```

```
    -2.9868  -1.2677  0.6304  -0.7009  1.1448  -2.6583  -0.3829  5.1866
    -------------------------
    Data Set 2 - Model 3
    -------------------------
    Error Rate = 0.6750
    σ1 = 5.855114052034532
    σ2 = 5.070305133828278


    ------------------------------------------------------------
    ------------------------------------------------------------
    Data Set 3
    ------------------------------------------------------------
    P(C1) = 0.2
    P(C2) = 0.8
    M1 = 0.2362, 1.3669, 0.3367, 0.5549, -0.0063, 0.9241, 0.3775, 0.4070
    M2 = 1.0253  0.4181  0.4563  0.3941  0.4005  0.6166  0.3171  0.9354
    -------------------------
    Data Set 3 - Model 1
    -------------------------
    Error Rate = 0.2400
    S1 =
    2.0559  -0.3032  -0.5313  -0.4405  1.3070  0.4624  -0.1037  -0.1044
    -0.3032  2.3104  -0.5280  -0.5306  -0.5244  -0.5516  -0.2588  0.2258
    -0.5313  -0.5280  3.1636  0.4187  -0.0829  -0.4318  -0.2902  0.2899
    -0.4405  -0.5306  0.4187  1.9929  -0.7272  -0.4296  0.8309  0.6381
    1.3070  -0.5244  -0.0829  -0.7272  2.6578  1.1431  -0.1203  -0.1389
    0.4624  -0.5516  -0.4318  -0.4296  1.1431  1.8975  -0.2714  -0.0160
    -0.1037  -0.2588  -0.2902  0.8309  -0.1203  -0.2714  1.2821  0.2571
    -0.1044  0.2258  0.2899  0.6381  -0.1389  -0.0160  0.2571  2.8011


    S2 =
    3.6564  0.1506  -0.0568  0.3361  -0.5260  0.5598  -0.1873  -0.2860
    0.1506  3.4698  -0.7366  0.0109  -0.0128  0.2459  -0.1359  -0.2992
    -0.0568  -0.7366  3.5129  -0.5629  -0.3901  0.2578  0.2308  0.5263
    0.3361  0.0109  -0.5629  3.9538  0.1582  0.2597  -1.4462  -0.2600
    -0.5260  -0.0128  -0.3901  0.1582  3.7715  -0.2777  -0.5899  -0.1331
    0.5598  0.2459  0.2578  0.2597  -0.2777  2.9816  0.0559  -0.6324
    -0.1873  -0.1359  0.2308  -1.4462  -0.5899  0.0559  5.3251  0.0595
    -0.2860  -0.2992  0.5263  -0.2600  -0.1331  -0.6324  0.0595  3.9993
    -------------------------
    Data Set 3 - Model 2
    -------------------------
    Error Rate = 0.2650
    S1 =
    2.8562  -0.0763  -0.2941  -0.0522  0.3905  0.5111  -0.1455  -0.1952
    -0.0763  2.8901  -0.6323  -0.2599  -0.2686  -0.1528  -0.1974  -0.0367
    -0.2941  -0.6323  3.3383  -0.0721  -0.2365  -0.0870  -0.0297  0.4081
    -0.0522  -0.2599  -0.0721  2.9733  -0.2845  -0.0849  -0.3077  0.1891
    0.3905  -0.2686  -0.2365  -0.2845  3.2146  0.4327  -0.3551  -0.1360
    0.5111  -0.1528  -0.0870  -0.0849  0.4327  2.4396  -0.1077  -0.3242
    -0.1455  -0.1974  -0.0297  -0.3077  -0.3551  -0.1077  3.3036  0.1583
    -0.1952  -0.0367  0.4081  0.1891  -0.1360  -0.3242  0.1583  3.4002


    S2 =
    2.8562  -0.0763  -0.2941  -0.0522  0.3905  0.5111  -0.1455  -0.1952
    -0.0763  2.8901  -0.6323  -0.2599  -0.2686  -0.1528  -0.1974  -0.0367
    -0.2941  -0.6323  3.3383  -0.0721  -0.2365  -0.0870  -0.0297  0.4081
    -0.0522  -0.2599  -0.0721  2.9733  -0.2845  -0.0849  -0.3077  0.1891
```

```
0.3905  -0.2686  -0.2365  -0.2845  3.2146  0.4327  -0.3551  -0.1360
0.5111  -0.1528  -0.0870  -0.0849  0.4327  2.4396  -0.1077  -0.3242
-0.1455  -0.1974  -0.0297  -0.3077  -0.3551  -0.1077  3.3036  0.1583
-0.1952  -0.0367  0.4081  0.1891  -0.1360  -0.3242  0.1583  3.4002
-------------------------
Data Set 3 - Model 3
-------------------------
Error Rate = 0.2150
σ1 = 2.1566489029880214
σ2 = 3.785871862079163
```

C

ERROR	Data Set 1	Data Set 2	Data Set 3
Method 1	0.290	0.045	0.240
Method 2	0.255	0.210	0.265
Method 3	0.650	0.675	0.215

This table shows that each set of data has a certain model that can be best repents the data. Meaning that there is no one method that can perfectly model any data set.

This distinction is very obvious in the table above. Where Data set 1 is best modeled by Method 2 and Data set 2 is best modeled by Method 1. Finally, Data set 1 is best modeled by Method 3. These are the reason for the difference in error rate in each data set:

- Model 2 implies that the covariance for class 1 and class 2 are close from each other.
- Model 1 implies that the covariance for class 1 and class 2 are not related to each other. That's leads to fewer assumption as it treats the data as dependent from each other with their unique covariance for each data.
- Model 3 implies that the data set is completely independent from each other as well as it treat is as it has the same variance for all dimensions.

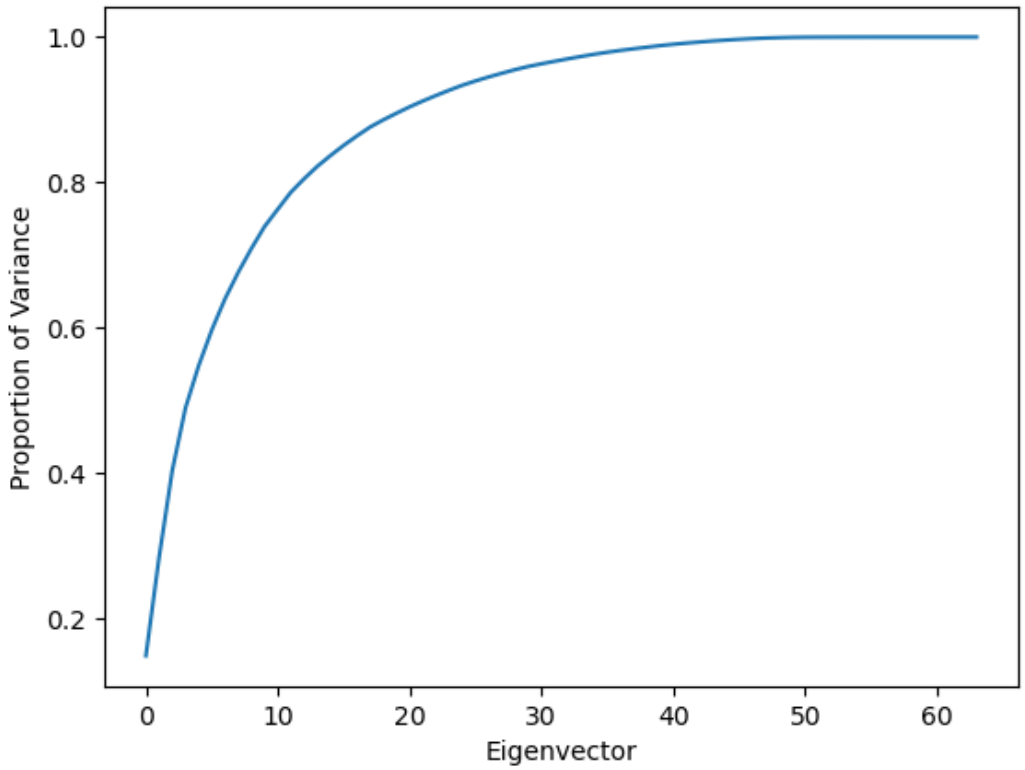
Problem 2:

A

k = 1 - Error Rate: 0.05387205387205387
k = 3 - Error Rate: 0.04713804713804714
k = 5 - Error Rate: 0.05387205387205387
k = 7 - Error Rate: 0.06060606060606061

B

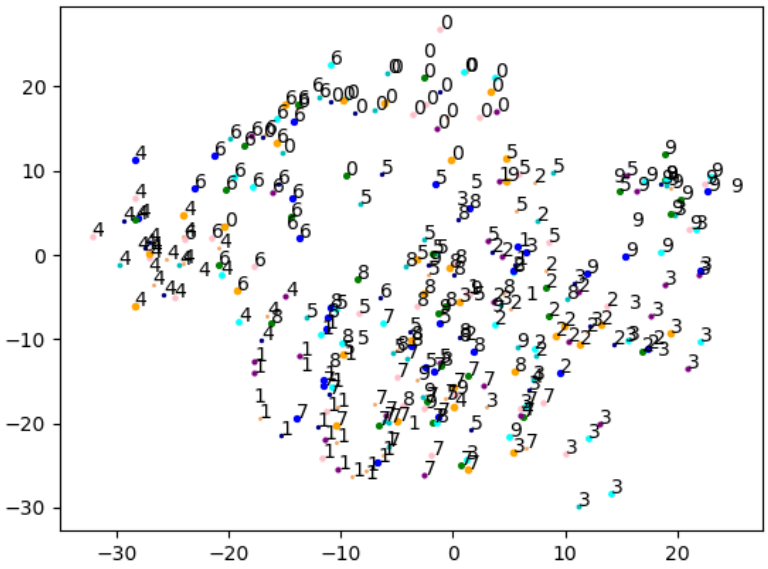
Minimum number of eigenvectors that explain at least 90% of the variance = 21



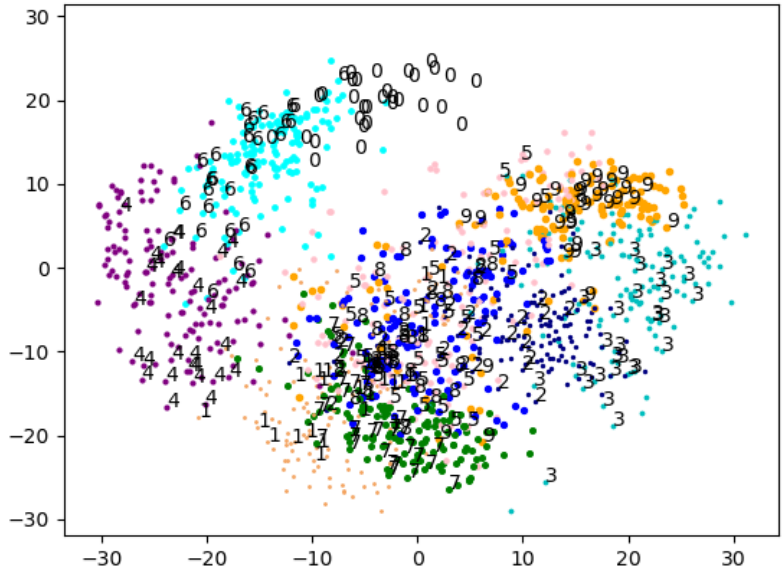
k = 1 - Error Rate: 0.04377104377104377
k = 3 - Error Rate: 0.037037037037037035
k = 5 - Error Rate: 0.04713804713804714
k = 7 - Error Rate: 0.05387205387205387

C

Training Set:



Test Set:

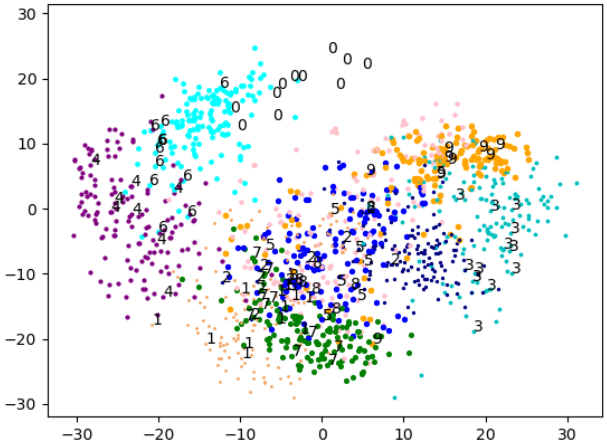


D

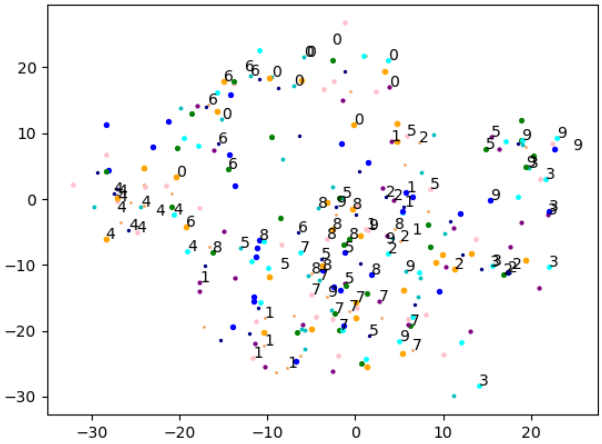
[L = 2 - k = 1] = Error Rate: 0.46464646464646464
[L = 2 - k = 3] = Error Rate: 0.4208754208754209
[L = 2 - k = 5] = Error Rate: 0.39057239057239057
[L = 4 - k = 1] = Error Rate: 0.1919191919191919
[L = 4 - k = 3] = Error Rate: 0.18518518518518517
[L = 4 - k = 5] = Error Rate: 0.1750841750841751
[L = 9 - k = 1] = Error Rate: 0.09764309764309764
[L = 9 - k = 3] = Error Rate: 0.09427609427609428
[L = 9 - k = 5] = Error Rate: 0.09764309764309764

E

Training Set:

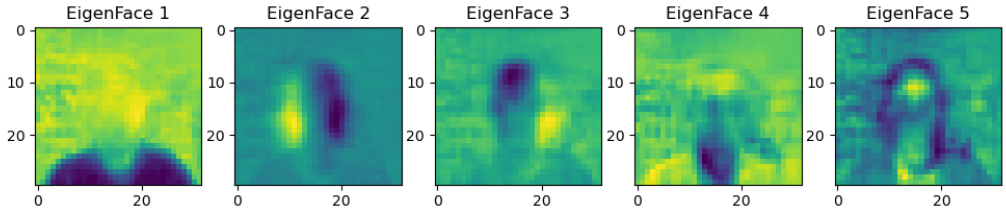


Test Set:



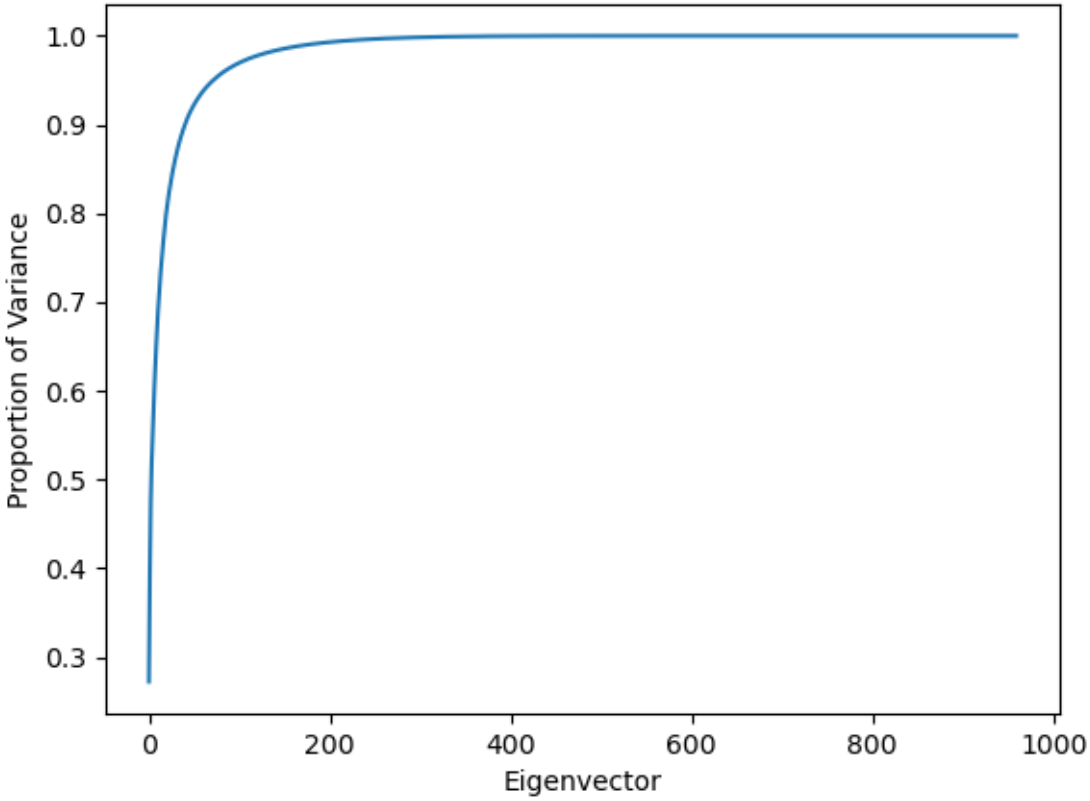
Problem 3:

A



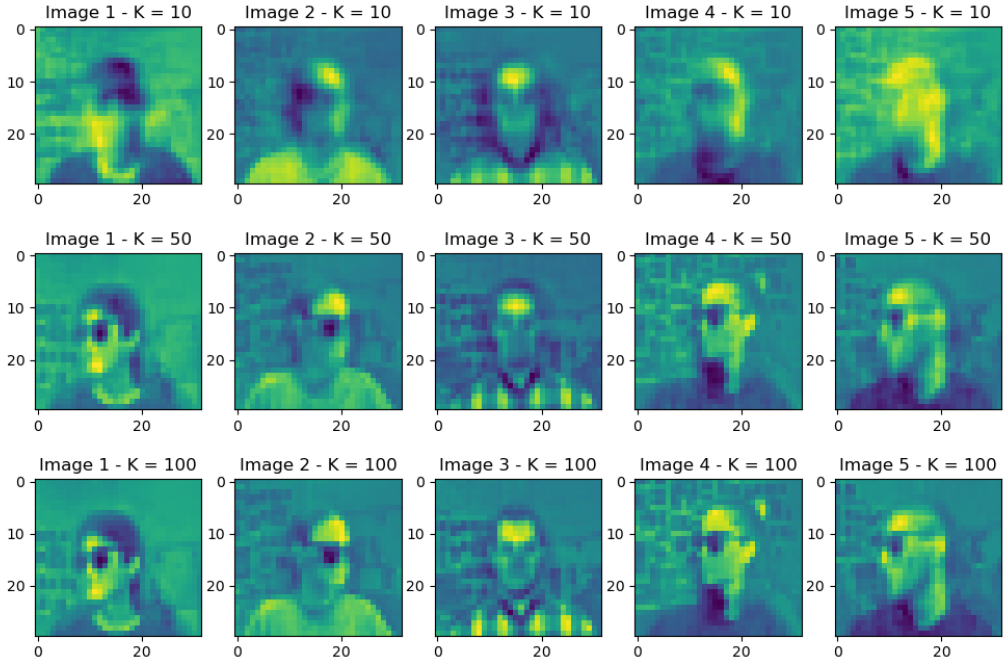
B

Minimum number of eigenvectors that explain at least 90% of the variance = **41**



k = 1 - Error Rate: 0.10483870967741936
k = 3 - Error Rate: 0.20161290322580644
k = 5 - Error Rate: 0.3225806451612903
k = 7 - Error Rate: 0.24193548387096775

C



We can see from the picture above that the larger k gets, the more details included in the picture. This is due to the increase in the information content as more principal components added.

This is being said, increasing principal components has a log increase in the information content. In this picture, 41 principal components will give a 90% or more details of original picture. It is seen in K=50 and k=100, there is little details, almost unnoticeable, added even though we doubled the principal component.