Korea Advanced Institute of Science and Technology

School of Electrical Engineering

EE488 Introduction to Machine Learning Spring 2018

Student’s Name: Dinh Vu

Student’s ID: 20184187

**Homework 3**

1. **PART 1**
2. **K-NN**
3. **Problem 1-i**

Because

The error rate can be express as:

1. **Problem 1-ii**

Bayes error rate is defined as below:

The sum of squares can be divided into 2 parts as follows:

According to Bunyakovsky for elements:

1. **Problem 1-iii**

Therefore, from :

1. **Problem 1-iv**

Because when the size of training data is infinite, the nearest neighbor of x will be itself, so:

Replace to :

Replace to :

Conclusion, the upper bound of the error rate in term of Bayes error rate is:

1. **Bayesian Statistic**
2. **Problem 2-i**

Prior:

Observe data , there are one-labels and 3 one-labels. The data size is .

1. **Problem 2-ii**

Prior , so:

According Bayes Rule:

Because given data from Gaussian distribution with mean and known variance :

Therefore:

Since is constant:

1. **Problem 2-iii**

According to Bayes Rule:

Since and are conditionally dependent variables:

is constant, so:

Therefore, from :

Finally:

1. **Problem 2-iv**

Exponential family:

Conjugate prior:

* Poisson distribution

Conjugate prior:

Where:

* Multinomial distribution

Conjugate prior:

Where:

* Laplace distribution

Conjugate prior:

Where:

* Dirichlet distribution

Conjugate prior:

* Gamma distribution

Conjugate prior (with known shape a):

Where: ,

**Table 1.1.** Summary of exponential family distributions and their conjugate prior

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Distribution** |  |  |  |  | **Conjugate prior** |
| Poisson |  | x |  |  | Gamma |
| Multinomial |  |  |  | 1 | Dirichlet |
| Laplace |  |  | 1 |  | Inverse Gamma |
| Dirichlet |  |  |  |  | Unknown |
| Gamma |  |  | 1 |  | Gamma |

1. **Decision Theory**

Expected loss or risk based on zero-one loss:

Where: is the major probable class.

* Case 1: . The major probable class is .

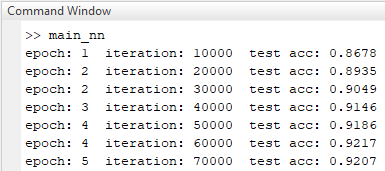
Expected loss or risk:

* Case 2: . The most probable class is .

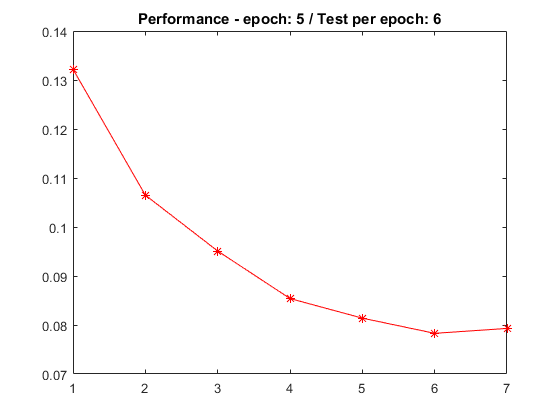
Expected loss or risk:

1. **Implementing Backpropagation Algorithm (Matlab Programming Assignment)**

Figure 4.1 and Figure 4.2 displays the test accuracy and the performance of each epoch. In the feed-forward process, the chosen activation function is sigmoid function. For the back-propagation, the derivative components are calculated in two for-loops which presents to layer. All weights of each layer are updated concurrently by matrix operations.



**Figure 4.2.** Test accuracy



**Figure 4.1.** Performance each epoch

1. **Recurrent Neural Network**
2. **Problem 5-i**

…

1. **Problem 5-ii**

Cross entropy:

Where:

1. **Problem 5-iii**

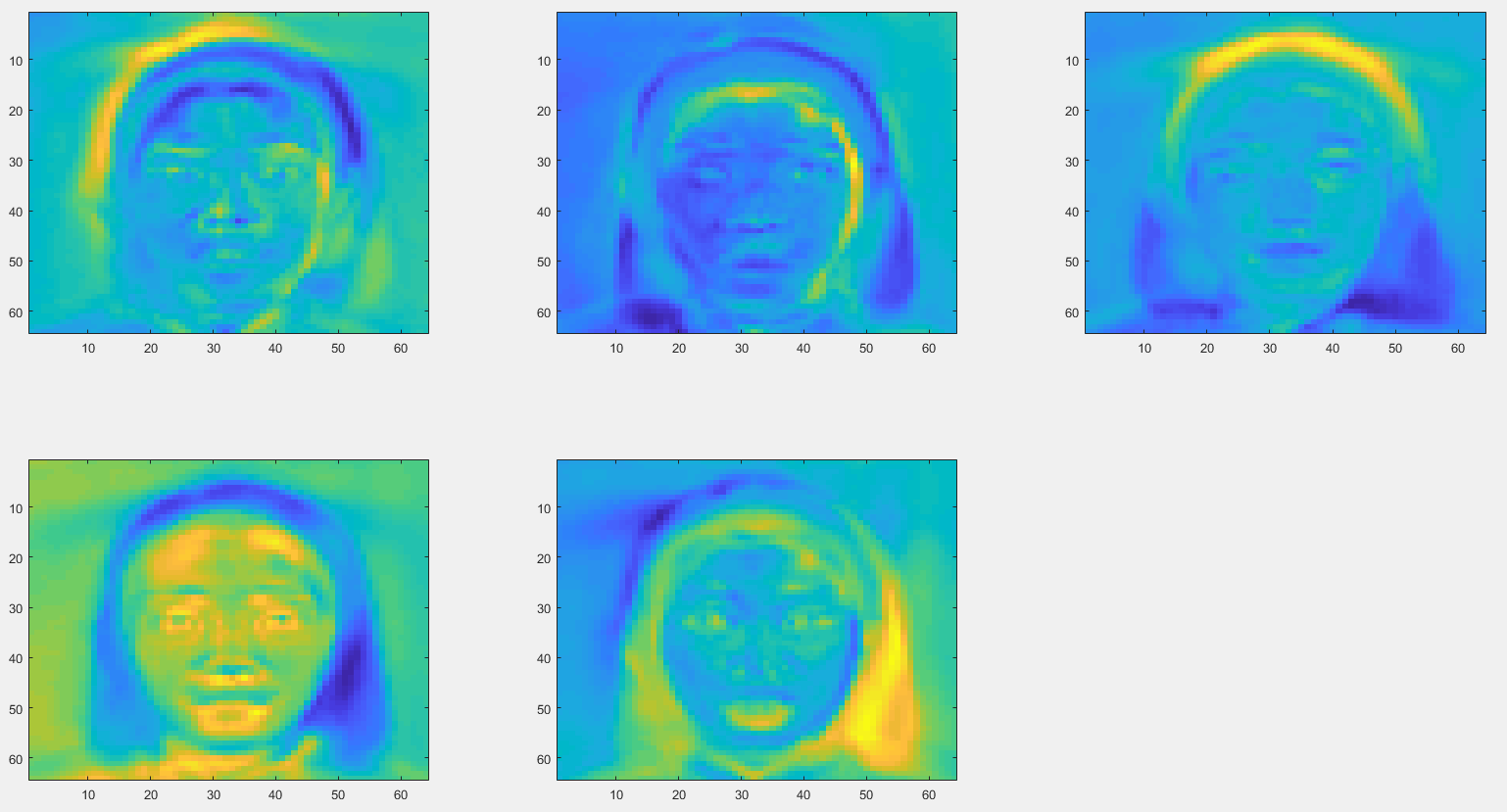
…

1. **PART 2**

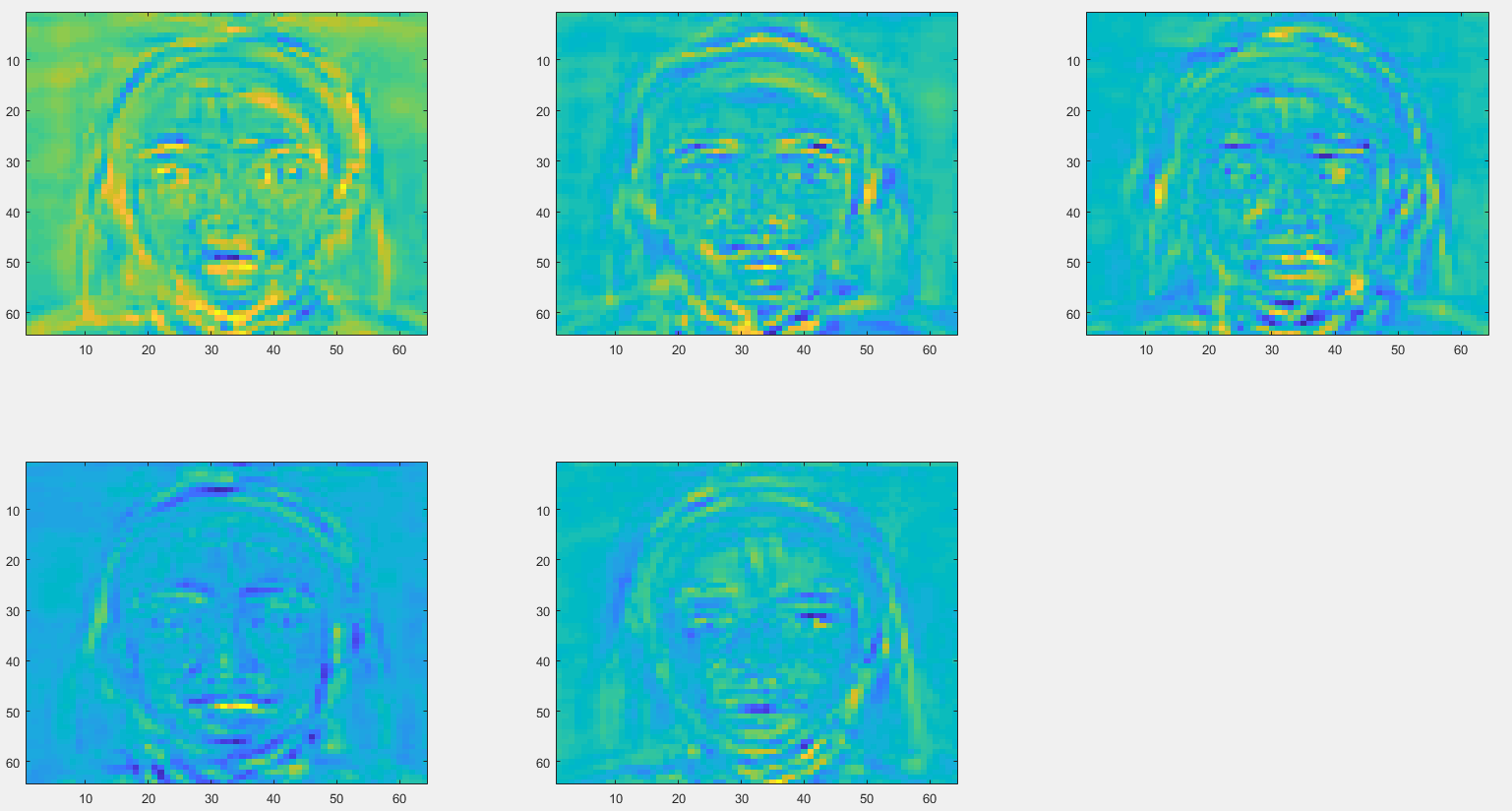
The results of PCA (Matlab Programming Assignment) are shown in Table 2.1 – 2.2 and Figure 2.1 – 2.4. Table 2.1 presents reconstruction error in each case k = 5, 50, 200 and 500. The top 5 principal components (PCs) are displayed from Figure 2.1 to Figure 2.4. Table 2.2 presents specifically the prediction about emotion of each test image in 4 cases of k.

**Table 2.1.** Reconstruction Error

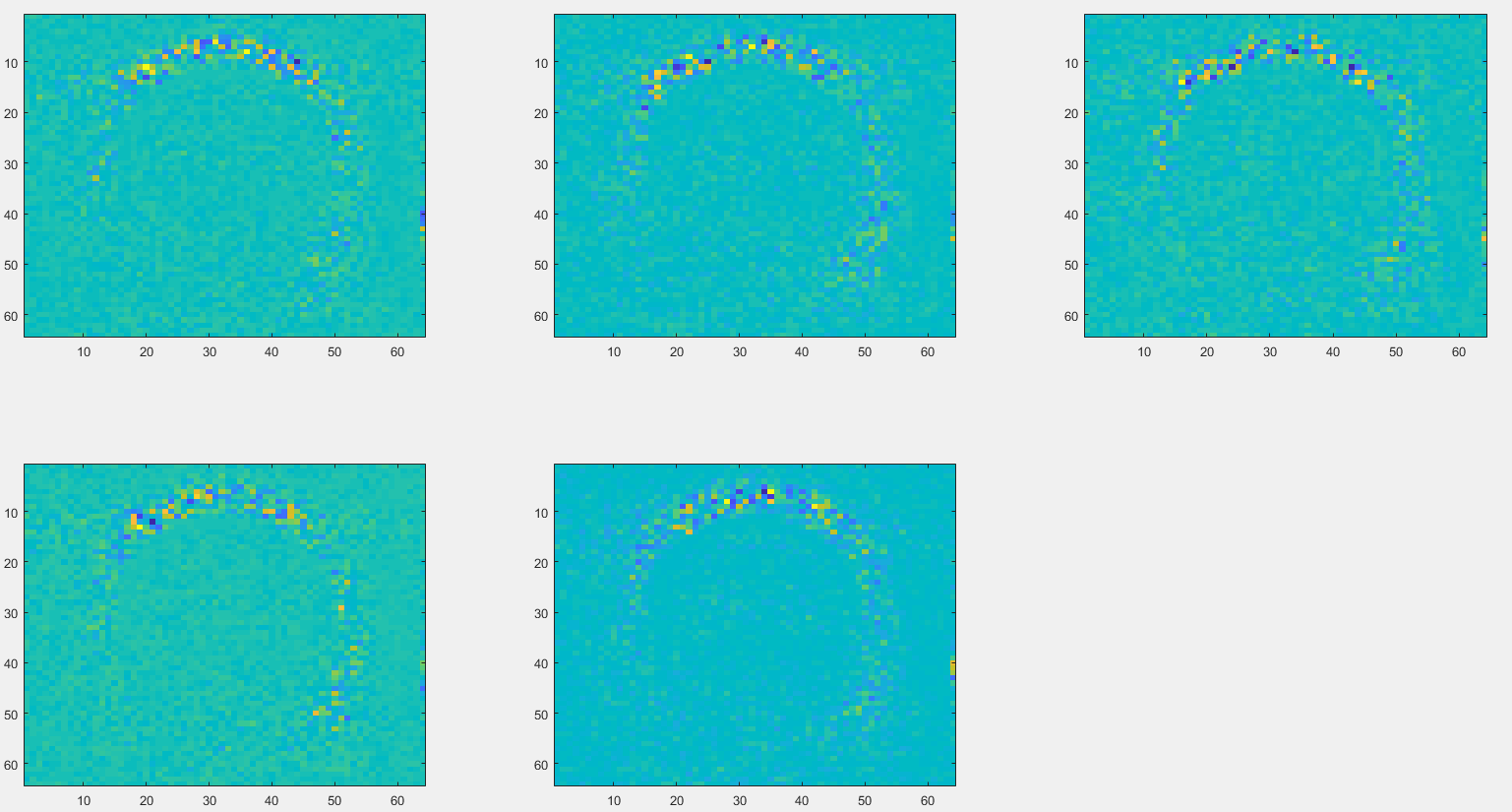
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **k** | 5 | 50 | 200 | 500 |
| **Reconstruction Error** | 2.1290 × 106 | 2.5866 × 105 | 4.0246 × 10-22 | 3.9890 × 10-22 |



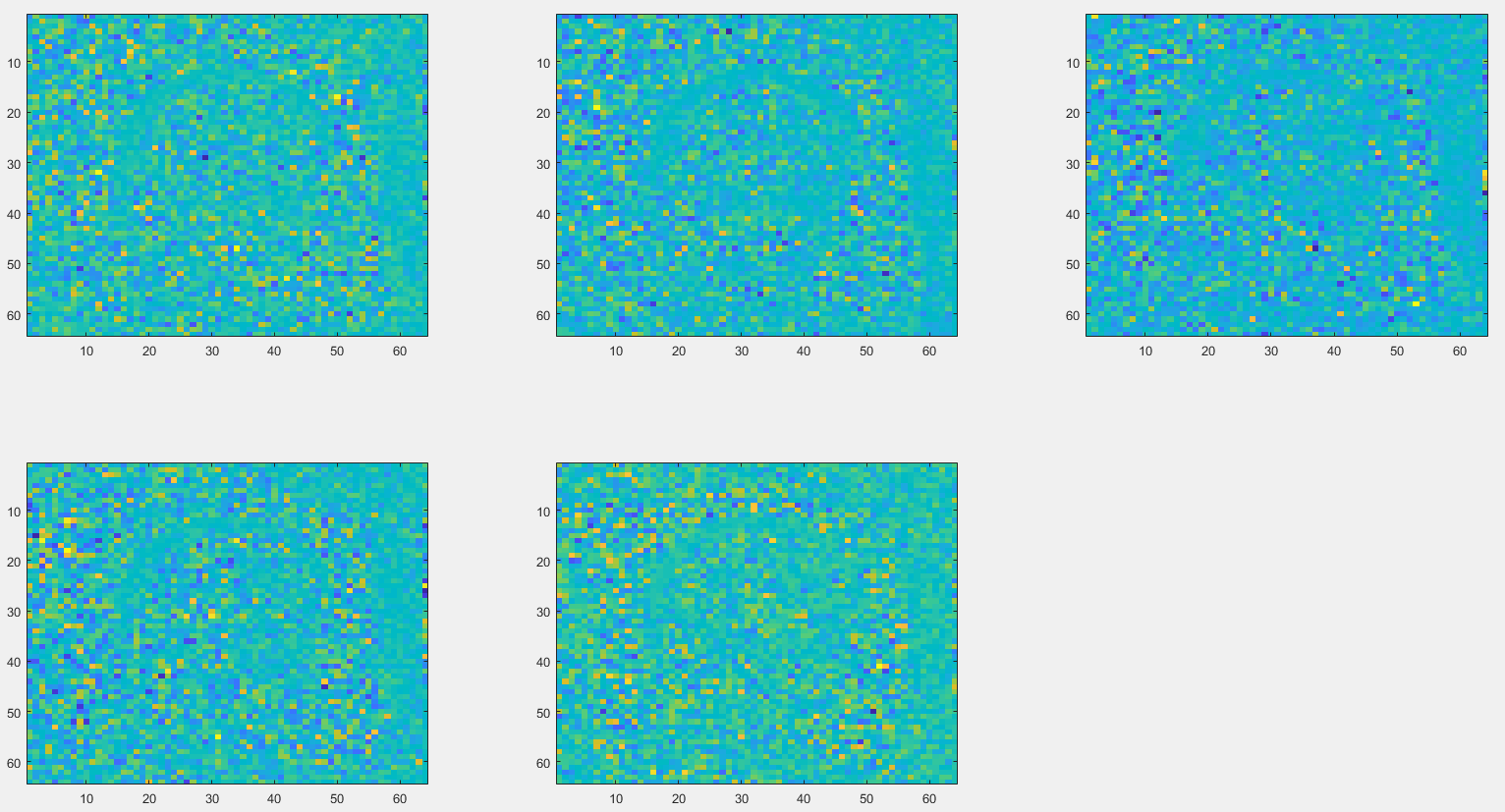
**Figure 2.1.** The top 5 PCs with k = 5



**Figure 2.2.** The top 5 PCs with k = 50



**Figure 2.3.** The top 5 PCs with k = 200



**Figure 2.4.** The top 5 PCs with k = 500

**Table 2.2.** The prediction results about emotion

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ID of test image** | **k** | | | | **Human eyes** | **Incorrect** |
| **5** | **50** | **200** | **500** |
| 1 | 7 | 18 | 18 | 18 | Not exist | All |
| 2 | 3 | 3 | 3 | 3 |  |  |
| 3 | 5 | 5 | 5 | 5 |  |  |
| 4 | 9 | 9 | 9 | 9 |  |  |
| **ID of test image** | **k** | | | | **Human eyes** | **Incorrect** |
| **5** | **50** | **200** | **500** |
| 5 | 14 | 14 | 14 | 14 |  |  |
| 6 | 19 | 19 | 19 | 19 |  |  |
| 7 | 22 | 21 | 22 | 22 | 21, 22 |  |
| 8 | 28 | 27 | 27 | 27 | 27, 28 |  |
| 9 | 30 | 30 | 30 | 30 |  |  |
| 10 | 35 | 35 | 35 | 35 |  |  |
| 11 | 36 | 36 | 36 | 36 |  |  |
| 12 | 44 | 44 | 44 | 44 |  |  |
| 13 | 46 | 46 | 46 | 46 |  |  |
| 14 | 48 | 49 | 49 | 49 | 49 | k = 5 |
| 15 | 52 | 52 | 52 | 52 |  |  |
| 16 | 55 | 55 | 55 | 55 |  |  |
| 17 | 56 | 55 | 55 | 55 | 55, 56 |  |
| 18 | 57 | 57 | 57 | 57 |  |  |
| 19 | 67 | 67 | 67 | 67 |  |  |
| 20 | 57 | 67 | 67 | 67 | 67 | k = 5 |
| 21 | 68 | 68 | 68 | 68 |  |  |
| 22 | 68 | 68 | 68 | 68 |  |  |
| 23 | 78 | 79 | 79 | 79 | 78, 79 |  |
| 24 | 77 | 81 | 81 | 81 | 81 | k = 5 |
| 25 | 73 | 83 | 83 | 83 | 83 | k = 5 |
| 26 | 85 | 85 | 85 | 85 |  |  |
| 27 | 94 | 94 | 94 | 94 | 92 | All |
| 28 | 94 | 92 | 92 | 92 | 92 | k = 5 |
| 29 | 93 | 94 | 94 | 94 | 93, 94 |  |
| 30 | 98 | 96 | 96 | 96 | 96 | k = 5 |
| 31 | 100 | 100 | 100 | 100 |  |  |
| 32 | 100 | 100 | 100 | 100 |  |  |
| 33 | 101 | 101 | 101 | 101 |  |  |
| 34 | 113 | 113 | 113 | 113 |  |  |
| 35 | 103 | 108 | 108 | 108 | 108 | k = 5 |
| 36 | 116 | 116 | 115 | 115 | 115 | k = 5, 50 |
| 37 | 118 | 118 | 118 | 118 |  |  |
| 38 | 119 | 119 | 119 | 119 | 119, 120 |  |
| 39 | 125 | 125 | 124 | 124 | 124, 125 |  |
| 40 | 129 | 129 | 129 | 129 |  |  |
| 41 | 137 | 134 | 134 | 134 | 134 | k = 5 |
| 42 | 144 | 140 | 140 | 140 | 140 |  |
| 43 | 144 | 144 | 144 | 144 |  |  |
| 44 | 144 | 144 | 144 | 144 |  |  |
| 45 | 145 | 145 | 145 | 145 |  |  |
| 46 | 148 | 148 | 148 | 148 |  |  |
| **ID of test image** | **k** | | | | **Human eyes** | **Incorrect** |
| **5** | **50** | **200** | **500** |
| 47 | 154 | 149 | 149 | 149 | 149 | k = 5 |
| 48 | 154 | 154 | 154 | 154 | 150 | All |
| 49 | 151 | 151 | 152 | 152 | 152 | k = 5, 50 |
| 50 | 154 | 154 | 154 | 154 |  |  |

Generally, all cases predict correctly about the face of person in the test images. However, the emotion prediction of the case k = 200 and 500 are the best. While the case k = 5 gives us the worst emotion prediction. There are 3 test images that all cases of k predicts emotion wrong and their ID equals to 1, 27 and 48. Conclusion, the PCA algorithm operates very well on face recognition but in emotion detection, it required improvement.