Bao Dang

CS 383

Homework 1

Part 1:

1. a.

p: samples in Class 1

n: samples in Class 2

Feature 1:

$p_{-8} = 1$, $n_{-8} = 0$

$p_{-5} = 1$, $n_{-5} = 0$

$p_{-3} = 1$, $n_{-3} = 0$

$p_{-2} = 1$, $n_{-2} = 1$

$p_{-1} = 0$, $n_{-1} = 1$

$p_0 = 1$, $n_0 = 0$

$p_1 = 0$, $n_1 = 1$

$p_5 = 0$, $n_5 = 1$

$p_6 = 0$, $n_6 = 1$

E(H(1)) = $4 \times \left(\frac{1+0}{5+5}\right) \times \left(\frac{-1}{1} \times \log_2\left(\frac{1}{1}\right) + -0 \times \log_2(0)\right) + 4 \times \left(\frac{0+1}{5+5}\right) \times \left(-0 \times \log_2(0) + \frac{-1}{1} \times \log_2\left(\frac{1}{1}\right)\right) + \left(\frac{1+1}{5+5}\right) \times \left(\frac{-1}{2} \times \log_2\left(\frac{1}{2}\right) + \frac{-1}{2} \times \log_2\left(\frac{1}{2}\right)\right) = 0.2$

IG(1) = $\left(\frac{-5}{10} \times \log_2\left(\frac{5}{10}\right) + \frac{-5}{10} \times \log_2\left(\frac{5}{10}\right)\right) - 0.2 = 0.8$

Feature 2:

$p_{-4} = 1$, $n_{-4} = 0$

$p_{-3} = 0$, $n_{-3} = 1$

$p_{-1} = 0$, $n_{-1} = 1$

$p_0 = 0$, $n_0 = 1$

$p_1 = 2$, $n_1 = 1$

$p_3 = 1$, $n_3 = 0$

$p_5 = 0, n_5 = 1$

$p_{11} = 1, n_{11} = 0$

E(H(2)) = $3 \times \left(\frac{1+0}{5+5}\right) \times \left(\frac{-1}{1} \times \log_2\left(\frac{1}{1}\right) + -0 \times \log_2(0)\right) + 4 \times \left(\frac{0+1}{5+5}\right) \times \left(-0 \times \log_2(0) + \frac{-1}{1} \times \log_2\left(\frac{1}{1}\right)\right) + \left(\frac{2+1}{5+5}\right) \times \left(\frac{-2}{3} \times \log_2\left(\frac{2}{3}\right) + \frac{-1}{3} \times \log_2\left(\frac{1}{3}\right)\right) = 0.2755$

IG(2) = $\left(\frac{-5}{10} \times \log_2\left(\frac{5}{10}\right) + \frac{-5}{10} \times \log_2\left(\frac{5}{10}\right)\right) - 0.2755 = 0.7245$

b. Feature 1 is more discriminating because IG(1) > IG(2)

2. LDA

$$\omega^T(\mu_1 - \mu_2)^T(\mu_1 - \mu_2)\omega - \lambda(\omega^T(\sigma_1^T\sigma_1 + \sigma_2^T\sigma_2)\omega) \qquad (1)$$

$$argmax_\omega(\omega^T(\mu_1 - \mu_2)^T(\mu_1 - \mu_2)\omega - \lambda(\omega^T(\sigma_1^T\sigma_1 + \sigma_2^T\sigma_2)\omega))$$

Taking the derivative and setting this equal to 0:

$$2(\mu_1 - \mu_2)^T(\mu_1 - \mu_2)\omega - \lambda\left((\sigma_1^T\sigma_1 + \sigma_2^T\sigma_2) + (\sigma_1^T\sigma_1 + \sigma_2^T\sigma_2)^T\omega\right) = 0$$

$$2(\mu_1 - \mu_2)^T(\mu_1 - \mu_2)\omega - \lambda\left((\sigma_1^T\sigma_1 + \sigma_2^T\sigma_2) + ((\sigma_1^T\sigma_1)^T + (\sigma_2^T\sigma_2)^T)\omega\right) = 0$$

$$2(\mu_1 - \mu_2)^T(\mu_1 - \mu_2)\omega - \lambda\left((\sigma_1^T\sigma_1 + \sigma_2^T\sigma_2) + (\sigma_1^T\sigma_1 + \sigma_2^T\sigma_2)\omega\right) = 0$$

$$2(\mu_1 - \mu_2)^T(\mu_1 - \mu_2)\omega - \lambda(2(\sigma_1^T\sigma_1 + \sigma_2^T\sigma_2)\omega) = 0$$

$$2(\mu_1 - \mu_2)^T(\mu_1 - \mu_2)\omega = \lambda(2(\sigma_1^T\sigma_1 + \sigma_2^T\sigma_2)\omega)$$

$$(\mu_1 - \mu_2)^T(\mu_1 - \mu_2)\omega = \lambda(\sigma_1^T\sigma_1 + \sigma_2^T\sigma_2)\omega$$

$$(\sigma_1^T\sigma_1 + \sigma_2^T\sigma_2)^{-1}((\mu_1 - \mu_2)^T(\mu_1 - \mu_2))\omega = \lambda\omega$$

Therefore, we must find the eigenvector/eigenvalue pairs, (ω, λ), for the equation above to get the maximum for equation (1)

Part 2:

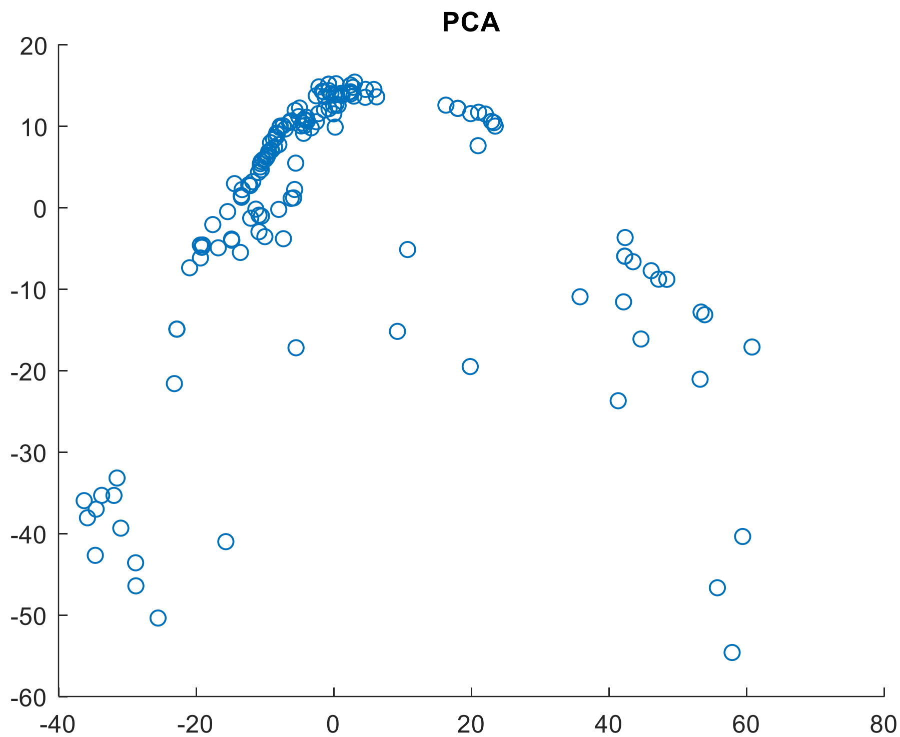

Figure 1. 2D PCA Projection of data

Part 3:

a) Number of principle components needed to represent 95% of information, k
   k = 37
b) Visualization of primary principle component

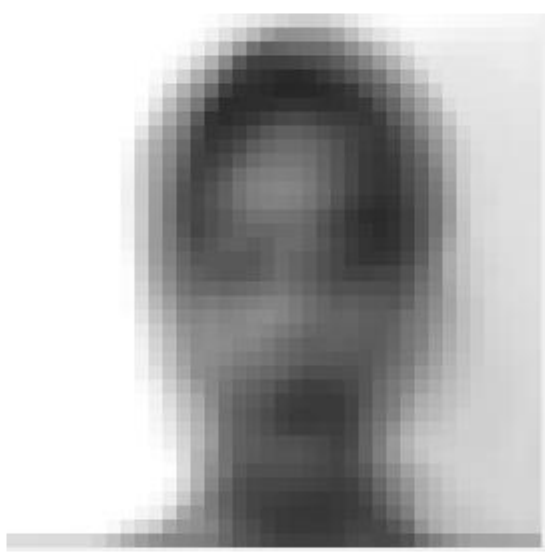

Figure 2. Primary Principle Component

c) Visualization of the reconstruction of the first person using the primary principle component and then using the k most significant eigen-vectors

Figure 3. Original image

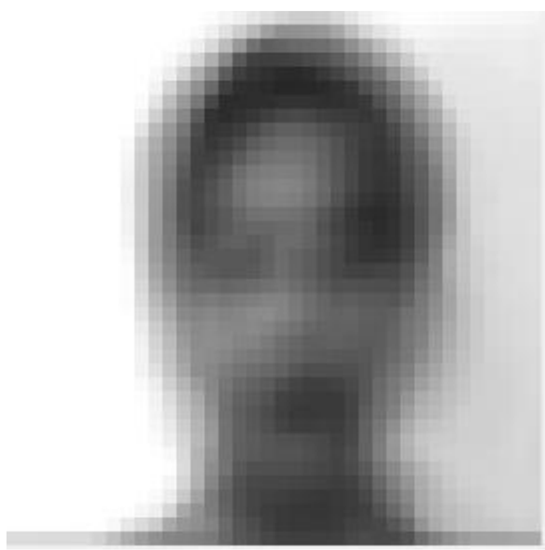

Figure 4. Single principle component

Figure 5. k principle component (k =37)