CSC411 A2

Guanxiong Liu
1002077726 liuguanx

1.

1. $P(x|\mu,\sigma) = \sum_k P(x|y=k,\mu,\sigma)\,P(y=k)$

$$P(y=k|x,\mu,\sigma) = \frac{P(x|y=k,\mu,\sigma)\cdot P(y=k)}{P(x|\mu,\sigma)}$$

$$= \frac{\alpha_k \exp\left\{-\sum_{i=1}^{D}\frac{1}{2\sigma_i^2}(x_i-\mu_{ki})^2\right\}\cdot\left(\prod_{i=1}^{D}2\pi\sigma_i^2\right)^{-1/2}}{\sum_k \alpha_k \exp\left\{-\sum_{i=1}^{D}\frac{1}{2\sigma_i^2}(x_i-\mu_{ki})^2\right\}\cdot\left(\prod_{i=1}^{D}2\pi\sigma_i^2\right)^{-1/2}}$$

$$= \frac{\alpha_k \exp\left\{-\sum_{i=1}^{D}\frac{1}{2\sigma_i^2}(x_i-\mu_{ki})^2\right\}}{\sum_k \alpha_k \exp\left\{-\sum_{i=1}^{D}\frac{1}{2\sigma_i^2}(x_i-\mu_{ki})^2\right\}}$$

2. $\ell(\theta;D) = -\log P(y^1,x^1,y^2,x^2,\ldots,y^N,x^N|\theta)$

$$= -\sum_{n=1}^{N}[y^n=k]\left(\sum_k \log\alpha_k - \frac{1}{2}\log\left(\prod_{i=1}^{D}2\pi\sigma_i^2\right) - \sum_{i=1}^{D}\frac{1}{2\sigma_i^2}(x_i^n-\mu_{ki})^2\right)$$

$$= -\sum_k \log\alpha_k \sum_{n=1}^{N}[y^n=k] + \frac{N}{2}\sum_{i=1}^{D}\log(2\pi\sigma_i^2) + \sum_k\sum_{i=1}^{D}\sum_{n=1}^{N}[y^n=k]\frac{1}{2\sigma_i^2}(x_i^n-\mu_{ki})^2$$

$$*: P(x,y) = \prod_k\left(\alpha_k \exp\left\{-\sum_{i=1}^{D}\frac{1}{2\sigma_i^2}(x_i-\mu_{ki})^2\right\}\right)^{[y=k]}$$

3. $\frac{\partial \ell(\theta;D)}{\partial \mu_{ki}} = \sum_{i=1}^{D}\sum_{n=1}^{N}\frac{1}{\sigma_i^2}[y^n=k]\,(x_i^n-\mu_{ki}) = 0$

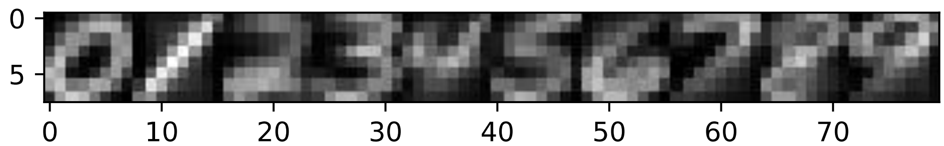
$$\mu_{ki} = \frac{\sum_{n=1}^{N}[y^n=k]x_i^n}{\sum_{n=1}^{N}[y^n=k]}$$

4. $\frac{\partial \ell(\theta;D)}{\partial \sigma_i^2} = \frac{N}{2}\,\frac{1}{\sigma_i^2} - \sum_k\sum_{n=1}^{N}[y^n=k](x_i^n-\mu_{ki})^2\cdot\frac{1}{2}\cdot\frac{1}{\sigma_i^4} = 0$

$$\frac{N}{2}\frac{1}{\sigma_i^2} = \sum_k\sum_{n=1}^{N}[y^n=k](x_i^n-\mu_{ki})^2\cdot\frac{1}{2}\cdot\frac{1}{\sigma_i^4}$$

$$\sigma_i^2 = \frac{\sum_k\sum_{n=1}^{N}[y^n=k](x_i^n-\mu_{ki})^2}{N}$$

CSC411 A2
Guanxiong Liu
1002077726 liuguanx
2.1.0

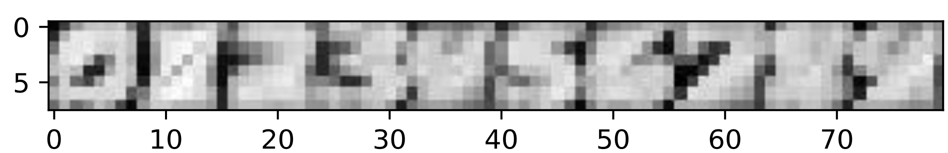


2.1.1
The accuracy with K=1 for Train Data = 100.0%
The accuracy with K=1 for Test Data = 96.875%
The accuracy with K=15 for Train Data = 96.37142857142858%
The accuracy with K=15 for Test Data = 96.1%

2.1.2
If we encounter a tie in our kNN classification. I choose the class of the closest train point in the candidate list to be the class of the test point. The candidate list consists of training points that have the highest frequency among k closest neighbors. For example, digit 0-9 have [1,1,1,1,1,1,2,2,3,3] frequency. Then digit 8 and 9 will be the candidate in my candidate list. Then I will find the closest point to the test point from the points of both class 8 and 9, and classify the test point to the class of closest point. I think this will be the best method to break a tie.

2.1.3
The optimal K=4
The average accuracy with optimal k = 96.55714285714285%
The accuracy with optimal k for Train Data = 98.64285714285714%
The accuracy with optimal k for Test Data = 97.275%

2.2.1



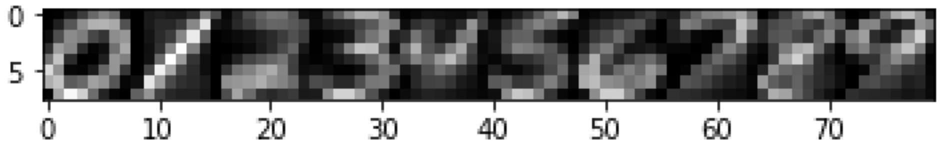
2.2.2
The average conditional log likelihood for training set is: -0.12462443666862986
The average conditional log likelihood for test set is: -0.1966732032552554

2.2.3
The accuracy on the training set is 98.14285714285714%
The accuracy on the test set is 97.275%

2.3.1

CSC411 A2
Guanxiong Liu
1002077726 liuguanx
Please see the code attached

2.3.2
Please see the code attached

2.3.3


2.3.4


2.3.5
The average conditional log likelihood for training set is: -0.9437538618002539
The average conditional log likelihood for test set is: -0.9872704337253583

2.3.6
The accuracy on the training set is 77.41428571428571%
The accuracy on the test set is 76.425%

2.4
KNN and Gaussian-conditionals perform the best with accuracy >97% on test set and >98% on the training set. Naïve Bayes performs the worst with accuracy about 76% on the test set and 77% on the training set. The results match my expectation since Naïve Bayes assume each feature (pixel in this case) is independent. We know in a picture, pixels are not independent, so Naïve Bayes is not appropriate for this classification. Therefore KNN and Gaussian-conditionals are more suitable models for this classification.