Malaviya National Institute of Technology, Surat
Department of Computer Science and Engineering

Pattern Recognition LAB Assignment, 2021

Day-1

1. Write MATLAB/Python program to generate the following distributions.

   - Normal distribution
   $$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1}$$
   - Uniform distribution
   $$p(x) = \frac{1}{b-a} \text{ if } a < x < b;\ 0 \text{ otherwise} \tag{2}$$
   - Exponential distribution
   $$p(x) = \lambda e^{-\lambda x} \text{ if } x \geq 0;\ 0 \text{ otherwise} \tag{3}$$
   - Poisson distribution
   $$p(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}; \lambda > 0 \tag{4}$$

2. Generate $N = 500$ 2-D random data points and plot its corresponding Gaussian PDF. $\left(p(x) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}\right)$

3. Given a set of d-dimensional samples. Write a program to find out covariance matrix. $\left(\Sigma = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \mu)(x_i - \mu)^T\right)$

Day-2

1. Generate $N = 500$ 2-D data points that are distributed according to the Gaussian distribution $N(m, \Sigma)$, with mean $m = [0, 0]^T$ and covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}$$

for the following cases

- $\sigma_1^2 = \sigma_2^2 = 1$, $\sigma_{12} = 0$
- $\sigma_1^2 = \sigma_2^2 = 0.2$, $\sigma_{12} = 0$
- $\sigma_1^2 = \sigma_2^2 = 2$, $\sigma_{12} = 0$
- $\sigma_1^2 = 0.2$, $\sigma_2^2 = 2$, $\sigma_{12} = 0$
- $\sigma_1^2 = 2$, $\sigma_2^2 = 0.2$, $\sigma_{12} = 0$
- $\sigma_1^2 = \sigma_2^2 = 1$, $\sigma_{12} = 0.5$
- $\sigma_1^2 = 0.3$, $\sigma_2^2 = 2$, $\sigma_{12} = 0.5$
- $\sigma_1^2 = 0.3$, $\sigma_2^2 = 2$, $\sigma_{12} = -0.5$

Plot each data set and comment the shape of the clusters formed by the data points.

2. Consider a c-class classification task in the d-dimensional space, where the data in all classes are distributed according to Gaussian distribution $N(m_i, S_i)$, $i = 1, 2, \ldots, c$.

For a given $m_i = [m_1, m_2, \ldots, m_d]$ and

$$S = \begin{bmatrix} \cdots & \cdots \\ \cdots & \cdots \end{bmatrix}_{d \times d}$$

Design a Bayesian classifier to classify a d-dimensional data into one of the c-classes. Assume $\sum P(w_i) = 1$, $i = 1, 2, \ldots, c$.

Day-3

1. For a given dataset (e.g., iris data set) D of size $N \times M$ with N: number of samples and M: number of features, design a Bayesian classifier to classify the test data. Divide the data set into training and testing data according to random percentage split.

2. For a given dataset (e.g., iris data set) D of size $N \times M$ with N: number of samples and M: number of features, design a (1) Euclidean distance classifier and (2) Mahalanobis distance classifier to classify the test data. Comment on the results. Assume classes are modeled by Gaussian distributions and classes to be equiprobable.

Day-4

1. Consider a dataset of your choice and implement k-means clustering algorithm.
2. Consider a dataset of your choice and implement agglomerative clustering algorithm.

——— xxx ———