EE 524 Machine Learning Lab

Assignment 4

17 October 2020

Please explain each line of your code. You can simply write a sentence for the same above the line or by the side of the line, describing it. This is to ensure you have understood the problem and formulated the code accordingly, and discourage any form of copying. Marks will be deducted as a penalty for copying.

Q.1. (a). Let us consider x to be a random variable, where x ∈ [-10,10]. Generate 100 samples of x and plot the pdf of x for x$\sim N(0,1)$ and x$\sim N(3,2)$. Comment on the difference in plot. If there is a difference in magnitude, comment why?

(b). Consider the two-dimensional normal distribution $p(x|w) \sim N(\mu, \Sigma)$ where $\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $\Sigma = \begin{bmatrix} 1 & 0.6 \\ 0.6 & 2 \end{bmatrix}$. Let $\mathbf{x}$ be a random vector defined by $\mathbf{x} = \begin{bmatrix} x1 \\ x2 \end{bmatrix}$, where $x_i \in [-10, 10]$. Plot the surface plot and contour plot of the pdf.

Q.2. Classify the iris flower dataset using a Naive Bayes Classifier. Link to download the dataset is here. The dataset consists of 3 classes and has 150 data samples. Each data sample has 4 features and a class label. The details related to the dataset can be found here.

Consider 70% of the data as training samples and the rest as testing samples. Display the predicted labels and the true labels. Also, calculate the accuracy of your prediction.

Note: 1. Arrange the samples according to class and calculate the prior probability of each class using a Gaussian distribution.

2. Don't arbitrarily choose 70% data for training. First arrange according to each class and then randomly choose the training and the testing samples from each class.

3. This is a supervised method. The testing labels should be used only to compare with the predicted labels. Also, you can opt for the confusion matrix(optional) to present your prediction. In case you are not using confusion matrix, show the predicted labels and true labels of each data sample using a bar diagram.