Week-6, Graded, Solution

Question-1

The prior probabilities for class c, where $c \in \{0,1,2,3\}$ is given by:

$$p_c = \frac{N_c}{\sum\limits_{r=0}^{3} N_r}$$

where, N_r is the number of points that belong to class r.

Question-2

The mean is given by the following equation. A slightly different notation is used here from the options, just to be consistent with the lectures:

$$\hat{\boldsymbol{\mu}}_{\boldsymbol{k}} = \frac{1}{N_k} \sum_{i=1}^{n} 1(y^{(i)} = k)\boldsymbol{x}^{(\boldsymbol{i})}$$

This equation means the following: get all the points that belong to class k. Take the mean of this set of points. This is the estimated mean for class k. Note that this mean is an $m-$dimensional vector, where m is the number of features. In some sense, this is nothing but the sample mean of points that belong to class k

The variance is given by:

$$\boldsymbol{v\hat{a}r_k} = \frac{1}{N_k} \sum_{i=1}^{n} 1(y^{(i)} = k)(\boldsymbol{x}^{(\boldsymbol{i})} - \hat{\boldsymbol{\mu}}_{\boldsymbol{k}}) \odot (\boldsymbol{x}^{(\boldsymbol{i})} - \hat{\boldsymbol{\mu}}_{\boldsymbol{k}})$$

Question-3

The correct mathematical expression for calculating the density function at x for the Gaussian distribution is:

$$\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

This translates to the following Python expression:

```
np.exp(-((x - mean) ** 2) / (2 * var)) / np.sqrt(2 * np.pi * var)
```

Question-4

We will take up just one data-point and see how to compute the posteriors. First, for class-1:

$$\begin{aligned}P(C_1 \mid x_1 = 0.25, x_2 = 3.14) &\propto P(C_1) \cdot P(x_1 = 0.25, x_2 = 3.14 \mid C_1)\\&= P(C_1) \cdot P(x_1 = 0.25 \mid C_1) \cdot P(x_2 = 3.14 \mid C_2)\\&= 0.5 \times 0.37 \times 0.55\\&= 0.102\end{aligned}$$

Now, for class-2:

$$\begin{aligned}P(C_2 \mid x_1 = 0.25, x_2 = 3.14) &\propto P(C_2) \cdot P(x_1 = 0.25, x_2 = 3.14 \mid C_2)\\&= P(C_2) \cdot P(x_1 = 0.25 \mid C_2) \cdot P(x_2 = 3.14 \mid C_2)\\&= 0.5 \times 0.63 \times 0.45\\&= 0.142\end{aligned}$$

From this we see that $(0.25, 3.14)$ can be predicted as class-2. The process is similar for other data-points. We have ignored the evidence (denominator) while computing the posterior as this going to be the same for both classes. That is why we have used the proportionality sign instead of the equality.

Question-5

The accuracy is given by:

$$\frac{TP + TN}{TP + TN + FP + FN} = \frac{39 + 42}{39 + 42 + 21 + 11} = 0.72$$

Question-6

For this question, we first look at all points that belong to each class and then take the mean of those points. For class-1, we have:

$$\mu_1 = \frac{[0.25, 3.14] + [0.75, 9.00] + [0.00, 10.00]}{3} = [0.33, 7.38]$$

A similar process can be applied to class-2.

Question-7

Now, we need to compute the standard deviation. First we compute the mean and then use it to compute the standard deviation. For class-1, the mean is:

$$\mu_1 = [0.49, 7.03]$$

For standard deviation, we compute some intermediate quantities:

x	$x - \mu_1$	$(x - \mu_1) \odot (x - \mu_1)$
$[0.25, 3.14]$	$[-0.24, -3.89]$	$[0.059, 15.16]$
$[0.48, 8.96]$	$[-0.01, 1.93]$	$[0.00017, 3.712]$
$[0.75, 9.00]$	$[0.26, 1.96]$	$[0.065, 3.867]$

Now, we can sum the third column and divide by 3 to get the variance. Square root of this would give the standard deviation:

$$\sigma_1 = [0.204, 2.753]$$

Note that all this computation was done on points belonging to class 1.