Algorithms for programs

Q4) for Bayesian classification we are

1. A loop will iterate over all values of K
2. We will randomly divide the dataset into k chunks and another loop will iterate over all k chunks
3. Particular i^{th} chunk = test and rest is training dataset
4. We are here using the Gaussian distribution for continuous dataset to find the probability of the occurrence of a given value to an attribute
   P(Y = 2 | attr_1, attr_2,.. attr_n) = P(Y = 2) * P(attr_1 | Y = 2) * P(attr_2 | Y = 2) * P(attr_3 | Y = 2)…. P(attr_n | Y = 2)
   similarly P(Y = 4 | attr_1, attr_4,.. attr_n) = P(Y = 4) * P(attr_1 | Y = 4) * P(attr_4 | Y = 4) * P(attr_3 | Y = 4)…. P(attr_n | Y = 4)
5. The final output of our test dataset will depend upon whichever probability is greater
6. In the end of the loop we will match with each of the answer of the test dataset and count the total number of error and hence calculate the accuracy

Q6)

a) Iterative re-weighted least squares. – As we know Equation for Iterative re-weighted for logistic regression is $$\mathbf{w}^{(\text{new})} = \mathbf{w}^{(\text{old})} - (\mathbf{\Phi}^{\mathrm{T}}\mathbf{R}\mathbf{\Phi})^{-1}\mathbf{\Phi}^{\mathrm{T}}(\mathbf{y}-\mathbf{t})$$
here y is the predicted output, t is the actual output and R is a diagonal matrix and
$$R_{nn} = y_n(1-y_n).$$
We will continue with 10 epochs

b) Stochastic Gradient Descent. : It works by taking alpha value and doing iteration over all training dataset and changing weight after computing hypothesis (predicted value) for each data point
$$\theta_j := \theta_j - \alpha(h_\theta(x^{(i)}) - y^{(i)})x_j^{(i)}$$
here θ_j is the j^{th} term of weight vector $h_\theta(x^{(i)})$ is the predicted value for data point x^i and $x_j^{(i)}$ is the j^{th} value of attribute of data point x^i and α is the learning rate