1. **Consider the following 3x3 confusion matrix (CM) summarising the testing results for a multi- class fruits dataset classification (27 marks in total).**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Confusion Matrix for Fruits dataset** | | **Predicted Class** | | |
| **Grapes** | **Apples** | **Melons** |
| **Actual Class** | **Grapes** | **10** | **4** | **1** |
| **Apples** | **2** | **15** | **1** |
| **Melons** | **1** | **1** | **13** |

* 1. **Provide the formula for calculating the sensitivity (or recall) in a CM (1 mark). Briefly explain what this index measures (2 mark).**
  2. **Provide the formula for calculating the specificity in a CM (1 mark). Briefly explain what this index measures (2 mark).**
  3. **What is the overall classification accuracy of this CM? (3 marks)**
  4. **Decompose the above 3x3 CM into three individual (per fruit) 2x2 CMs (3x2marks= 6 marks)**
  5. **What is the sensitivity and specificity for each class? (6x2marks = 12 marks) Show all steps of your work.**

1. **Suppose, we have the following three numbers: 10, 15 and 20, and we wish to normalise them via the z-score method. Show all steps of your work for this normalisation process (3 x 2 marks**

**= 6 marks).**

1. **Explain the differences between supervised and unsupervised learning processes (2 marks). Provide an example of machine algorithm for each of these two processes (2 marks).**
2. **Summarise the strength(s) (2 marks) and the weakness(es) (2 marks) of K-means clustering.**
3. **We have two clusters, X, and *Y*, containing a total of four points, where each point is defined by an integer value in one dimension, X = {10} and Y = {2, 4, 6}. You need to calculate the distance of these two clusters based on three different styles: single (2 marks), complete (2 marks) and average linkage (2 marks). Show all steps of your work (6 marks in total).**