Government Engineering College, Modasa



**INDEX**

**Machine Learning(3710216)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sr.**  **No.** | **Date** | **Title** | **Pg. No.** | **Sign** |
| 1. |  | Given the following vectors:  A = [1, 2, 3, 4, 5, 6, 7, 8, 9 10]  B = [4, 8, 12, 16, 20, 24, 28, 32, 36, 40]  C = [10, 9, 8, 7, 6, 5, 4, 3, 2, 1]  Ex. 1: Find the arithmetic mean of vector A, B and C  Ex. 2: Find the variance of the vector A, B and C  Ex. 3: Find the euclidean distance between vector A and B  Ex. 4: Find the correlation between vectors A & B and A & C |  |  |
| 2. |  | Load breast cancer dataset and perform classification using Euclidean distance. Use 70% data as training and 30% for testing. |  |  |
| 3. |  | Repeat the above experiment with 10-fold cross validation and find the standard deviation in accuracy. Repeat the above experiment with 10-fold cross validation and find the standard deviation in accuracy. |  |  |
| 4. |  | Repeat the experiment 2 and build the confusion matrix. Also derive Precision, Recall and Specificity of the algorithm |  |  |
| 5. |  | Predict the class for X = < Sunny, Cool, High, Strong > using Naïve Bayes Classifier for given data…. |  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 6. |  | For the data given in Exercise 5, find the splitting attribute at first level: |  |  |
| 7. |  | Generate and test decision tree for the dataset in exercise 5 |  |  |
| 8. |  | Find the clusters for following data with k = 2:  Start with points 1 and 4 as two separate clusters. |  |  |
| 9. |  | Find following statistics for the data given in Exercise 1 |  |  |
| 10. |  | Given the following vectors:  X = [340, 230, 405, 325, 280, 195, 265, 300, 350,  310]; %sale  Y = [71, 65, 83, 74, 67, 56, 57, 78, 84, 65];  Ex. 1: Find the Linear Regression model for independent variable X and dependent variable Y.  Ex. 2: Predict the value of y for x = 250. Also find  the residual for y4. |  |  |