**FR. CONCEICAO RODRIGUES COLLEGE OF ENGG**.
Fr. Agnel Ashram, Bandstand, Bandra (W) Mumbai 400 050.

**Machine Learning (ML-CSC701)**
**2023-2024**
**II UNIT TEST Question Bank**

<table>
<tr><th>Sr. No</th><th>Questions</th></tr>
<tr><td>1</td><td>
Apply the DBSCAN algorithm with similarity threshold of 0.85 (using the similarity matrix ) to the given data points and MinPts>=2 (Minimum required points in a cluster) what are core, border and noise (outliers) in the set of points given in table.
<table>
<tr><td></td><td>P1</td><td>P2</td><td>P3</td><td>P4</td><td>P5</td></tr>
<tr><td>P1</td><td>1.00</td><td>0.30</td><td>0.60</td><td>0.70</td><td>0.79</td></tr>
<tr><td>P2</td><td>0.30</td><td>1.00</td><td>0.88</td><td>0.90</td><td>0.89</td></tr>
<tr><td>P3</td><td>0.60</td><td>0.88</td><td>1.00</td><td>0.60</td><td>0.90</td></tr>
<tr><td>P4</td><td>0.70</td><td>0.90</td><td>0.60</td><td>1.00</td><td>0.80</td></tr>
<tr><td>P5</td><td>0.79</td><td>0.89</td><td>0.90</td><td>0.80</td><td>1.00</td></tr>
</table>
</td></tr>
<tr><td>2</td><td>
The cluster has the following data point. Calculate their intra cluster distance. Also calculate its inertia .
<table>
<tr><th>X1</th><th>X2</th><th>Class</th></tr>
<tr><td>182</td><td>72</td><td>2</td></tr>
<tr><td>170</td><td>56</td><td>1</td></tr>
<tr><td>168</td><td>60</td><td>1</td></tr>
<tr><td>179</td><td>68</td><td>2</td></tr>
<tr><td>185</td><td>72</td><td>2</td></tr>
<tr><td>188</td><td>77</td><td>2</td></tr>
</table>
</td></tr>
<tr><td>3</td><td>
Apply the Xgboost on the following dataset:-<br>
Draw the tree for only one sequence. Fill the values in the following table.
<table>
<tr><th>X1</th><th>X2</th><th>X3</th><th>Residual-1</th><th>Output of Tree_1</th><th>Residual-2</th></tr>
<tr><td>1</td><td>1</td><td>30K</td><td></td><td></td><td></td></tr>
<tr><td>1.5</td><td>1</td><td>31k</td><td></td><td></td><td></td></tr>
<tr><td>2.5</td><td>0</td><td>41k</td><td></td><td></td><td></td></tr>
<tr><td>3</td><td>0</td><td>50k</td><td></td><td></td><td></td></tr>
<tr><td>3.5</td><td>1</td><td>52k</td><td></td><td></td><td></td></tr>
</table>
</td></tr>
<tr><td>4</td><td>
Plot the hyperplane for the following points:-<br>
(2,2),(3,2),(2,-2),(3,-2),(5,0),(6,2)(6,-2),(7,0). Also Classify the point(6,2) based on the calculated hyperplane.<br><br>
Note:- Ensure that you follow all the steps accurately, including the creation of graphs when necessary. Display the classified point by substituting values into the hyperplane equation and also depict it on the graph.
</td></tr>
</table>

| | |
|---|---|
| 5 | **Calculate EigenValue and EigenVector for the following centered features**<br><br>CentreX1 / CentreX2<br>1, 0<br>1, 2<br>3, 4<br>-3, -4<br>-1, -2<br>-1, 0 |
| 6 | **Construct the RBF pattern classifier for the following dataset:-**<br><br>A / B / Y<br>0, 0, 0<br>0, 1, 1<br>1, 0, 1<br>1, 1, 0 |
| 7 | **Calculate the PCA Values using the following Normalized Eigenvectors on non centered features:-**<br><br>X1 / X2<br>130, 78<br>128, 80<br>128, 82<br>126, 78<br>128, 80<br>128, 82<br><br>$v_1 = \begin{bmatrix} 0.59 \\ 0.81 \end{bmatrix} \quad \lambda_1 = 12.08$<br><br>$v_2 = \begin{bmatrix} -0.81 \\ 0.59 \end{bmatrix} \quad \lambda_2 = 0.32$ |
| 8 | Why do we need similarity or dissimilarity measure for clustering algorithms? |
| 9 | Clustering algorithms are sensitive to starting points. Justify the statement with example. Also, explain impact of outliers, distance measure and noise on K-mean clustering algorithm. |
| 10 | Why we need to use dimensionality reduction methods? PCA minimizes loss of information. Justify. |
| 11 | What is bagging and boosting? How it is used to combine classifiers to improve results? |
| 12 | Explain soft margin and hard margin SVM with suitable example. |
| 13 | What is a kernel? How you will choose appropriate kernel for problem in hand? |
| 14 | Compute the principal component of following data-<br><br>CLASS 1<br>X = 2, 3, 4<br>Y = 1, 5, 3<br><br>CLASS 2 |

| | |
|---|---|
| | X = 5, 6, 7<br>Y = 6, 7, 8 |
| 15 | Find appropriate transformation to convert non-linearly separable data to linearly separable data:<br>Positive class:<br>$\left\{\begin{pmatrix}2\\2\end{pmatrix},\begin{pmatrix}2\\-2\end{pmatrix},\begin{pmatrix}-2\\-2\end{pmatrix},\begin{pmatrix}-2\\2\end{pmatrix}\right.$<br>Negative Class:<br>$\left\{\begin{pmatrix}1\\1\end{pmatrix},\begin{pmatrix}1\\-1\end{pmatrix},\begin{pmatrix}-1\\-1\end{pmatrix},\begin{pmatrix}-1\\1\end{pmatrix}\right\}$ |