Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

SLOT: F2+T

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - II
FALL SEMESTER 2025-2026

Programme Name & Branch	**: B.Tech CSE (Specialization in Data Science)**
Course Code and Course Name	**: BCSE334L - Predictive Analytics**
Faculty Name(s)	**: Dr. ANBARASI M, Dr. SANJIBAN SEKHAR ROY, Dr. GUNAV Dr. SAYAN SIKDER, Dr. THANGARAMYA K, Dr. PADMAVA**
Class Number(s)	**: VL2025260101654/1649/1652/1657/1660/4**
Date of Examination	**: 10.10.2025**
Exam Duration	**: 90 minutes** **Maximum Mark**

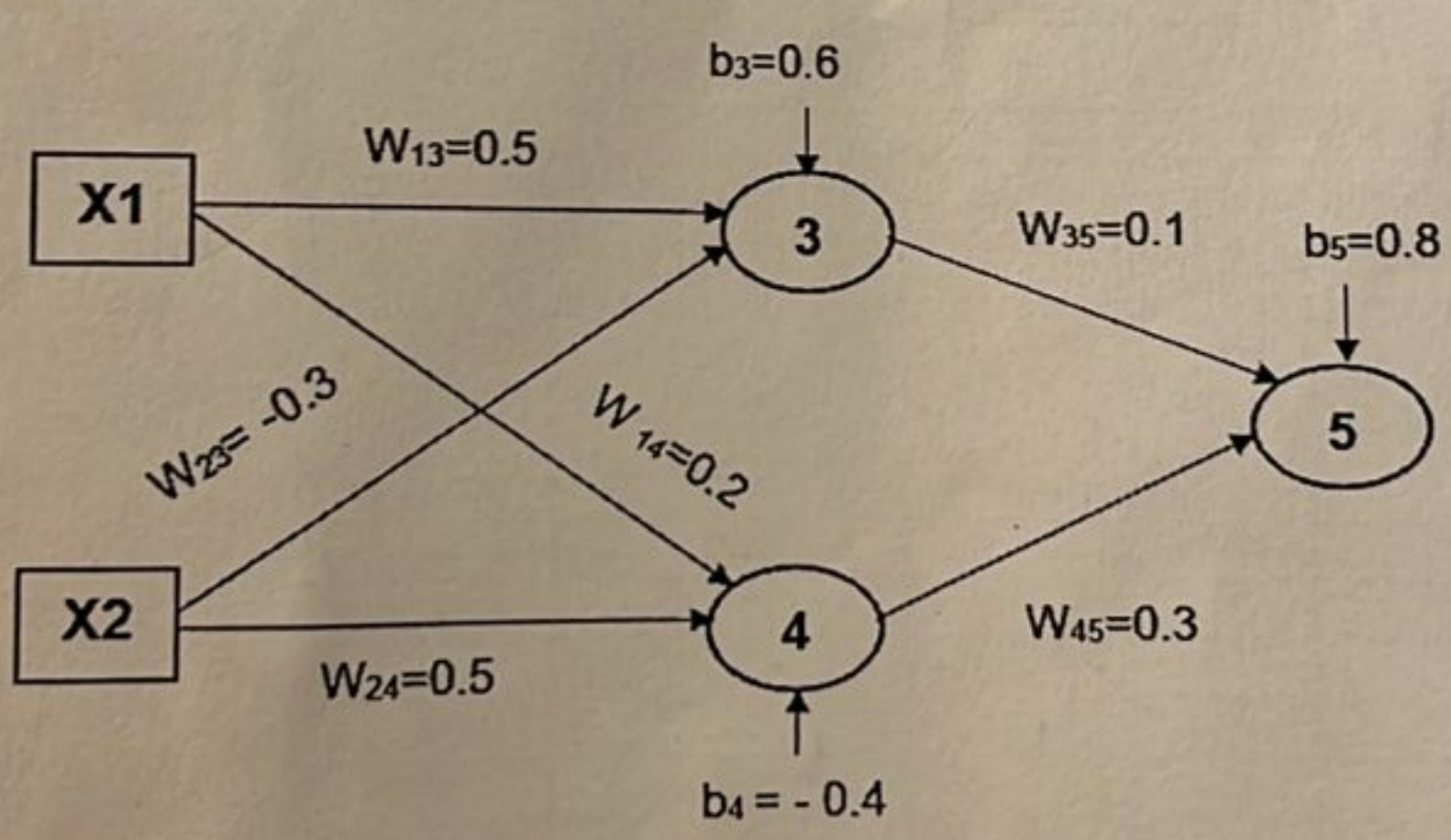
General instruction(s):

- Answer All Questions. M - Max mark; CO – Course Outcome; BL – Blooms Ta Level (1 – Remember, 2 – Understand, 3 – Apply, 4 – Analyse, 5 – Evaluate, 6 – Crea
- Course Outcomes: CO3: Apply regression and classification model on applica decision making and evaluate the performance.

<table>
<tr><th>Q. No</th><th>Question</th><th>Module</th></tr>
<tr><td>1.</td><td>Construct the quadratic regression model and find the regression equation for the following data. Calculate the value of Y when X = 7.
<table>
<tr><td>X</td><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td>Y</td><td>1</td><td>4</td><td>9</td><td>15</td></tr>
</table></td><td>4</td></tr>
<tr><td>2.</td><td>(i) The table given below represents the Height, Age values and the corresponding Weight of 10 individuals. Find the Weight of the 11th record using K-Nearest Neighbor regression algorithm with 2 different K values (K=3 and K=5).
<table>
<tr><th>ID</th><th>Height</th><th>Age</th><th>Weight</th></tr>
<tr><td>1</td><td>5</td><td>45</td><td>77</td></tr>
<tr><td>2</td><td>5.11</td><td>26</td><td>47</td></tr>
<tr><td>3</td><td>5.6</td><td>30</td><td>55</td></tr>
<tr><td>4</td><td>5.9</td><td>34</td><td>59</td></tr>
<tr><td>5</td><td>4.8</td><td>40</td><td>72</td></tr>
<tr><td>6</td><td>5.8</td><td>36</td><td>60</td></tr>
<tr><td>7</td><td>5.3</td><td>19</td><td>40</td></tr>
<tr><td>8</td><td>5.8</td><td>28</td><td>60</td></tr>
<tr><td>9</td><td>5.5</td><td>23</td><td>45</td></tr>
<tr><td>10</td><td>5.6</td><td>32</td><td>58</td></tr>
<tr><td>11</td><td>5.5</td><td>38</td><td>?</td></tr>
</table>
(ii) How do we handle a kNN regression problem when the independent features/predictors are categorical in nature? Give examples.</td><td>4</td></tr>
<tr><td>3</td><td>Consider the following dataset which includes demographic information and credit scores of individuals. Apply the Naive Bayes classification to predict the credit score of a new individual based on the following demographic features: Age>40, Gender=Male, Income=Medium, Education=Graduate, Occupation=Skilled.</td><td>5</td></tr>
</table>

Age	Gender	Income	Education	Occupation	Credit Score
<40	Male	High	Graduate	Professional	Low
<40	Female	Medium	High School	Managerial	Low
>40	Male	High	Postgraduate	Professional	High
>40	Female	High	Postgraduate	Professional	High
>40	Male	Low	High School	Skilled	Low
<40	Female	Medium	Graduate	Managerial	Medium
<40	Male	Medium	Graduate	Professional	High
>40	Female	High	Postgraduate	Managerial	High
>40	Male	High	High School	Skilled	Low
<40	Female	Medium	Graduate	Professional	Medium
>40	Male	Medium	High School	Skilled	Medium
>40	Female	Low	Graduate	Skilled	High

4. Consider the neural network given below with initial values of weights and biases. The learning rate is considered as 0.5. Train the network for the training tuple (1, 1, 0), where last number is the target output. Show weight and bias updates after one forward and one backward propagation. Assume that sigmoid activation function is used in the network. 5



5. (i) The actual marks scored by a student in five different subjects are [50, 60, 70, 80, 90]. The regression algorithm predicted these values as [48, 65, 68, 75, 100]. Calculate the Mean Absolute Error, Mean Squared Error and Root Mean Squared Error. 4

(ii) A machine learning model is trained to predict tumour in patients. 1349 people who have tumours are predicted positively by the model. 147 people who don't have tumours are predicted negatively by the model. 54 people are predicted as positive of having a tumour, although they don't have a tumour. 1 person who has tumours is predicted as negative. Find F1-Score of the model. 5