JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -2 EXAMINATION- 2026

B.Tech-VI Semester (CSE)

COURSE CODE (CREDITS): 18B1WCI635 (2) MAX. MARKS: 25

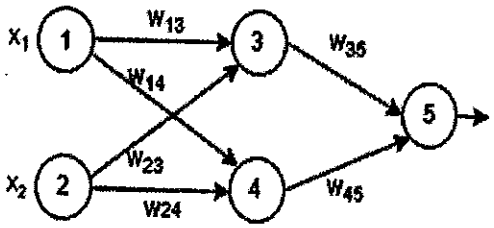
COURSE NAME: Data Mining and Data Warehousing

COURSE INSTRUCTORS: EGA, MGN, PTK MAX. TIME: 1.5 Hrs.

Note: *(a) All questions are compulsory.*
(b) The candidate is allowed to make Suitable numeric assumptions wherever required for solving problems
(c) Calculator is allowed.

JUIT TEST-2 EXAMINATION-March-2026

Q. No.	Question	CO	Marks
Q1.	An online retail company maintains a data warehouse to analyze its business performance. The warehouse stores information related to sales transactions, product details, customer information, time of purchase, and store location hierarchy (city, state, country). The company also wants to organize the location and product information into multiple related tables to reduce data redundancy while supporting analytical queries such as sales by region and product category performance.	2	
	a) Identify the most appropriate data warehouse schema for this scenario and justify your answer with valid reason.		[2]
	b) Design the schema structure by identifying the fact table and the related dimension tables.		[2]
	c) Explain one advantage and one disadvantage of using this schema in the given scenario.		[1]
Q2.	Consider the following Artificial Neural Network trained using the backpropagation algorithm. The learning rate is 0.1. The initial weights, biases and the target output are given in the table below. Using the sigmoid activation function, find the updated values of $w35, w45, \theta_3$, and θ_4 after one iteration of backpropagation.	4	[5]



Weight/bias	Values	Weight/bias	Values
$x1$	1	$w35$	0.1
$x2$	1	$w45$	0.3
$w13$	0.5	θ_3	0.6
$w14$	0.2	θ_4	-0.4
$w23$	-0.3	θ_5	0.8
$w24$	0.5	T	0

<table>
<tr><td>Q3.</td><td>A supermarket chain records daily billing, product purchases, and payment details in a database where thousands of transactions occur every minute. The system must process each transaction quickly and accurately.
a) Identify the type of system used to process these transactions. Justify your answer.
b) Write any four key differences between Online Transaction Processing and Online Analytical Processing systems.</td><td>2</td><td>[1]
[2]</td></tr>
<tr><td>Q4.</td><td>Given the following training dataset, which predicts whether a patient has a disease (Yes/No) based on their Exercise Level (High, Medium, Low) and Junk Food Habit (Yes/No), apply the C4.5 algorithm to determine the root node of the decision tree.
<table>
<tr><th></th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th></tr>
<tr><td>Exercise Level</td><td>Low</td><td>Low</td><td>Medium</td><td>Medium</td><td>High</td><td>High</td></tr>
<tr><td>Junk Food Habit</td><td>Yes</td><td>No</td><td>Yes</td><td>No</td><td>Yes</td><td>No</td></tr>
<tr><td>Disease</td><td>Yes</td><td>Yes</td><td>Yes</td><td>No</td><td>No</td><td>No</td></tr>
</table></td><td>4</td><td>[5]</td></tr>
<tr><td>Q5.</td><td>Consider the following dataset. Apply the K-Nearest Neighbors algorithm with K = 3 to predict the price of a car having engine power of 130 HP and age of 3 years. Use L-1 norm as the distance metric. Perform any necessary preparation of the features before computing distances.
<table>
<tr><th>Engine Power (HP)</th><th>Age (years)</th><th>Price (₹ lakh)</th></tr>
<tr><td>80</td><td>1</td><td>4</td></tr>
<tr><td>100</td><td>2</td><td>5</td></tr>
<tr><td>120</td><td>3</td><td>6</td></tr>
<tr><td>140</td><td>4</td><td>7</td></tr>
<tr><td>160</td><td>5</td><td>8</td></tr>
</table></td><td>4</td><td>[5]</td></tr>
<tr><td>Q6.</td><td>Explain the concept of margin in a Support Vector Machine (SVM). Also, justify why a linear SVM is referred to as a maximal margin classifier.</td><td>4</td><td>[2]</td></tr>
</table>

JUIT TEST-2 EXAMINATION-April-2026