JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -1 EXAMINATION- 2025

B.Tech. - VI Semester (IT/MIT)

COURSE CODE (CREDITS): 18B11CI613 (3) MAX. MARKS: 15

COURSE NAME: Data Mining

COURSE INSTRUCTORS: Ekta Gandotra MAX. TIME: 1 Hour

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.

<table>
<tr><th>Q. No.</th><th>Question</th><th>CO</th><th>Marks</th></tr>
<tr><td>Q1.</td><td>a. Analyze the key characteristics of a data warehouse and evaluate how it supports the KDD process in extracting valuable insights for decision-making.
b. Consider the data: 25, 30, 45, 50, 60. Apply Min-Max normalization method on this data to scale the values between -1 and 1.</td><td>CO1</td><td>[3]
[2]</td></tr>
<tr><td>Q2.</td><td>a. Consider two text documents represented as term frequency vectors in a 6-dimensional space:
Doc1 = (2,3,0,5,7,1), Doc2 = (4,1,2,6,3,2)
Compute the cosine similarity of these two documents. Also interpret the results.
b. Consider the daily step count (in thousands) of two fitness enthusiasts over five days:
Person A = (7, 10, 12, 9, 11), Person B = (8, 9, 14, 10, 12) Compute the Manhattan distance between their step counts. Interpret the results in terms of the similarity of their physical activity levels.</td><td>CO3</td><td>[2]
[2]</td></tr>
<tr><td>Q3.</td><td>a. How do you determine the value of k in the k-NN algorithm? What are the drawbacks of choosing a value of k that is too small?
b. Given the dataset below, apply ID3 algorithm to determine the root node for the decision tree.
<table>
<tr><th>Age</th><th>Income</th><th>Buy Product?</th></tr>
<tr><td>Young</td><td>High</td><td>No</td></tr>
<tr><td>Young</td><td>Low</td><td>Yes</td></tr>
<tr><td>Middle</td><td>High</td><td>Yes</td></tr>
<tr><td>Middle</td><td>Low</td><td>Yes</td></tr>
<tr><td>Old</td><td>High</td><td>Yes</td></tr>
<tr><td>Old</td><td>Low</td><td>No</td></tr>
</table></td><td>CO4</td><td>[2]
[4]</td></tr>
</table>