JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -2 EXAMINATION- 2025

M.Tech-I Semester (CSE/IT)

COURSE CODE (CREDITS): 22M1WCI131 (3) MAX. MARKS: 25

COURSE NAME: Data Warehousing and Data Mining

COURSE INSTRUCTORS: Dr. Rakesh Kanji MAX. TIME: 1 Hour 30 Min

Note: *(a) All questions are compulsory.*

(b) The candidate is allowed to make Suitable numeric assumptions wherever required for solving problems

<table>
<tr><th>Q.No</th><th>Question</th><th>CO</th><th>Marks</th></tr>
<tr><td>Q1</td><td>Explain the idea of Decision tree to classify the data. Describe its significance to deal with large number of attributes with a suitable diagram. Justify the use of Shanon entropy formula in ID3 or c4.5 algorithm.</td><td>3,4</td><td>1+2+2</td></tr>
<tr><td>Q2</td><td>
<table>
<tr><th>Day</th><th>Outlook</th><th>Temperature</th><th>Humidity</th><th>Play</th></tr>
<tr><td>1</td><td>Sunny</td><td>Hot</td><td>High</td><td>Yes</td></tr>
<tr><td>2</td><td>Sunny</td><td>Hot</td><td>High</td><td>No</td></tr>
<tr><td>3</td><td>Overcast</td><td>Cold</td><td>Low</td><td>Yes</td></tr>
<tr><td>4</td><td>Overcast</td><td>Cold</td><td>Low</td><td>No</td></tr>
<tr><td>5</td><td>Overcast</td><td>Cold</td><td>Medium</td><td>Yes</td></tr>
<tr><td>6</td><td>Overcast</td><td>Cold</td><td>Medium</td><td>No</td></tr>
</table>
Apply suitable algorithm between ID3 or c4.5 to create the decision tree with 2 level only.</td><td>3,4</td><td>5</td></tr>
<tr><td>Q3</td><td>Proof that objective function for linear regression is convex. Show the derivation of gradient descent technique and identify its significance to deal non convex function,</td><td>4</td><td>2+2+1</td></tr>
<tr><td>Q4</td><td>Apply suitable encoding on to above table and find the independent parameters (β) for logistic regression model. Show pictorially the strength and weakness of this model to deal with classification.</td><td>3,4</td><td>3+2</td></tr>
<tr><td>Q5</td><td>Explain the idea of Support vector machine to deal with limitation of logistic regression model and outlier. Derive the SVM objective function.</td><td>3,4</td><td>2+3</td></tr>
</table>