Pardeep Kumar

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -2 EXAMINATION- October 2017

M.Tech III Semester

COURSE CODE: 13M1WCI331 MAX. MARKS: 25

COURSE NAME: MACHINE LEARNING

COURSE CREDITS: 03 MAX. TIME: 1Hr 30 Min

Note: *All questions are compulsory. Carrying of mobile phone during examinations will be treated as case of unfair means.*

1. Consider the weather data set given below:

Outlook	Temperature	Humidity	Windy	Play
Sunny	Hot	High	False	No
Sunny	Hot	High	True	No
Overcast	Hot	High	False	Yes
Rainy	Mild	High	False	Yes
Rainy	Cool	Normal	False	Yes
Rainy	Cool	Normal	True	No
Overcast	Cool	Normal	True	Yes
Sunny	Mild	High	False	No
Sunny	Cool	Normal	False	Yes
Rainy	Mild	Normal	False	Yes
Sunny	Mild	Normal	True	Yes
Overcast	Mild	High	True	Yes
Overcast	Hot	Normal	False	Yes
Rainy	Mild	High	True	No

Table 1: Weather Data Set

Make a classification model using ID-3 decision tree approach to predict whether the game will be played or not. [12]

2. Suppose we are interested in analyzing transactions at ALLELECTRONICS with respect to the purchase of computer games and videos. Of the 10000 transactions analyzed, the data show that 6000 of the customer transactions included computer games, while 7500 included videos and 4000 included both computer games and videos. (i) Make the contingency table and calculate degree of freedom (ii) Compute the correlation using chi-square statistics. [7]
3. (i) "CART is better than C4.5 and ID-3 decision trees" (ii) " Entropy signifies the impurity in datasets" Justify the above statements with suitable reasoning, example and diagram. [6]