JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -1 EXAMINATION- Feb 2020

B.Tech. (IT) VI Semester

COURSE CODE: 10B22CI622   MAX. MARKS: 15

COURSE NAME: Data Mining

COURSE CREDITS: 4   MAX. TIME: 1 Hr.

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

Q1.   CO-1   Describe whether or not each of the following jobs is a data mining task.   [3]
     (a) Dividing the customers of a company according to their profitability.
     (b) Predicting the outcomes of tossing a (fair) pair of dice.
     (c) Predicting the future stock price of a company using historical records.

Q2.   CO-1   Classify the following attributes as binary, discrete, or continuous. Also classify them as   [3] qualitative (nominal or ordinal) or quantitative (interval or ratio). Some cases may have more than one interpretation, so briefly indicate your reasoning if you think there may be some ambiguity.
     (i) Time in terms of AM or PM. (ii) Brightness as measured by a light meter
     (iii) Height above sea level (iv) ISBN numbers for books

Q3.   CO-1   Suppose that the data for analysis includes the attribute age. The age values for the data tuples are **35, 35, 35, 35, 36, 13, 15, 16, 20, 21, 22, -5, 22, 25, 25, 25, 25, 30, 33, 33, 40, 45, 46, 52, -9, 70, 16, 19, 20.**
     (a) Identify and remove if outliers are present   [1]
     (b) What is the mean of the data? What is the median?   [1]
     (c) What is the mode of the data? Comment on the data's modality   [1]
     (d) Plot an equal-width histogram of width 10.   [2]

Q4.   CO-2   What is Entity Identification Problem in Data preprocessing? Propose a methods for   [2] handling this problem

Q5.   CO-2   Differentiate between the following.   [2]
     (a) Spatial data and spatiotemporal data.
     (b) Time-series data and temporal data.