SPRING MID SEMESTER EXAMINATION-2019
School of Computer Engineering
KALINGA INSTITUTE OF INDUSTRIAL TECHNOLOGY
DEEMED TO BE UNIVERSITY, BHUBANESWAR-24

DATA WAREHOUSING AND DATA MINING
[CS-6301]

Time: 1½ Hours **Full Mark: 20**

Answer any four questions including question No.1 which is compulsory.
The figures in the margin indicate full marks. Candidates are required to give their answers in their own words as far as practicable and all parts of a question should be answered at one place only.

Q.1. [5x1]
(a) Define data mining and discuss its importance in information technology.
(b) What condition makes association rules are interesting?
(c) Define anti-monotone property.
(d) What is the need of information gain in decision tree classifier?
(e) What condition two item sets A and B will have "No Correlation between them"?

Q.2.
(a) What is Interquartile range (IQR)? Draw an IQR with an odd sample size as ***80, 74,87, 73, 91, 82, 86, 74, 91.*** [2]
(b) Calculate both Eigen vectors & Eigen values for $A = \begin{pmatrix} 2 & -1 \\ 5 & -2 \end{pmatrix}$ [3]

Q.3.
(a) Define Support and Confidence in the context of frequent pattern mining. [2]
(b) What is the Apriori property? Using Apriori Algorithm find the final item set for the following dataset S. [3]
(*Where minimum support is 0.5*)

S=

TID	Items
101	Book, Note, Pen
102	Pencil, Note, Eraser
103	Book, Pencil, Note, Eraser
104	Pencil, Eraser

Q.4.
(a) Explain information gain as an attribute selection measure in decision tree. [2]
(b) Consider the following database of houses represented by 5 training examples. The target attribute is 'Acceptable', which can have values 'yes' or 'no'. This is to be predicted based on the other attributes of the house. [3]

House	Furniture	Nos. rooms	New kitchen	Acceptable
1	No	3	Yes	Yes
2	Yes	3	No	No
3	No	4	No	Yes
4	No	3	No	No
5	Yes	4	No	Yes

Construct the decision tree from the given examples, that would be learned by the ID3 algorithm.

Q.5.
(a) What is Bessel's correction? Calculate and Draw a box plot for the data set ***5, 5, 8, 9, 6, 3, 3, 12, 2, 6, 2, 7, 3, 9, 11.*** [2]
(b) The ages of the 112 people who live on a tropical island are grouped as follows, Calculate Estimated Mean, Estimated Median, & Estimated Mode of living people. [3]

Age	Number
0-9	20
10-19	21
20-29	23
30-39	16
40-49	11
50-59	10
60-69	7
70-79	3
80-89	1

****