



AUTUMN MID SEMESTER EXAMINATION-2019
School of Computer Engineering
KALINGA INSTITUTE OF INDUSTRIAL TECHNOLOGY
DEEMED TO BE UNIVERSITY, BHUBANESWAR-24
DATA MINING AND DATA WAREHOUSING
[CS-6301/CS-6342]

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) How is a data warehouse different from a database?
- (b) Define anti-monotone property.
- (c) What is Posterior Probability?
- (d) What condition two item sets A and B will have “No Correlation between them”?
- (e) Differentiate between classification and prediction in data mining.

Q.2.

- (a) With a neat diagram explain the architecture of data mining. [3]

- (b) How are they different from traditional data warehouse? [2]

Q.3. What is the Apriori property? Using Apriori Algorithm find the final item set for the following dataset S. [5]

S=

TID	Items
101	Milk,Bread,Butter
102	Jam,Bread,Egg
103	Milk,Jam,Bread,Egg
104	Jam,Egg

Q.4. Measure your dependent/correlated events LIFT computation for the following contingency table, [5]

play basketball ⇒ *eat cereal* [40%, 66.7%],

play basketball ⇒ *not eat cereal* [20%, 33.3%]

	Basketball	Not basketball	Sum (row)
Cereal	2000	1750	3750
Not cereal	1000	250	1250
Sum(col.)	3000	2000	5000

Q.5. What is Posterior Probability? Use Bayes Theorem to find out “*If a patient has fever, what’s the probability he/she has Typhoid?*” [5]

- (i). Normally, Typhoid causes high fever 60% of the time.
- (ii). Prior probability of any patient having Typhoid is 5/10000,
- (iii). Prior probability of any patient having fever is 1/50
