

Total No. of Questions : 6]

SEAT No. :

P5807

[Total No. of Pages : 2

BE/Insem/Oct.-585
B.E. (Computer Engineering)
Data Analytics (Theory)
(2015 Pattern) (Semester - I)

Time : 1 Hour]

[Max. Marks : 30

Instructions to the candidates:

- 1) *Solve Q.1 or Q.2, Q.3 or Q.4, Q.5 or Q.6.*
- 2) *Figures to the right indicate full marks.*
- 3) *Assume suitable data if necessary.*

Q1) a) What is Big data? Explain characteristics of Big Data. [4]

b) Explain different phases of data analytics life cycle. [6]

OR

Q2) a) Explain current analytic architecture with suitable diagram. [6]

b) Explain Bigdata Ecosystem. [4]

Q3) a) What is clustering? Explain k-means clustering algorithm with use cases. [6]

b) Explain Hypothesis testing with example. [4]

OR

Q4) a) Explain any two of the following : [4]

- i) Wilcoxon rank-sum test
- ii) Type 1 and Type 2 errors
- iii) ANOVA

P.T.O.

b) Height Weight

[6]

185	72
170	56
168	60
179	68
182	72
188	77
180	71
180	70
183	84
180	88
180	67
177	76

Use the above data and group them using k-means clustering algorithm. Show calculation of centroids.

Q5) a) What is market basket analysis? Explain Apriori algorithm with example. [6]

b) Explain logistic regression. Explain use cases of logistic regression. [4]

OR

Q6) a) Transactional Data for an All Electronics Branch is as follows : [6]

TID	List of Item _ IDs
T100	I1, I2, I5
T200	I2, I4
T300	I2, I3
T400	I1, I2, I4
T500	I1, I3
T600	I2, I3
T700	I1, I3
T800	I1, I2, I3, I5
T900	I1, I2, I3

Find the frequent item set and generate association rules with confidence values.

b) What is regression? Explain any one type of regression in detail. [4]

