[6144]-306
S.Y.B.B.A. (Computer Application)
CA - 305 : BIG DATA
(2019 CBCS Pattern) (Semester -III)

Time : 2½ Hours]   *[Max. Marks : 70*
Instructions to the candidates:
   1)   All questions are compulsory.
   2)   Figures to right indicates full marks.

Q1) Attempt any EIGHT of the following: **[8×2=16]**

   a)   What is predictive analytics?

   b)   What are the applications of Association Rule Mining?

   c)   What is WEKA?

   d)   Write any two advantages of Big Data.

   e)   Define Regression Analysis.

   f)   Write four applications of Data Science.

   g)   Write any two needs of Machine Learning.

   h)   Write two disadvantages of EM algorithm.

   i)   What is R?

   j)   Enlist any four tools used for Big Data.

Q2) Attempt any FOUR of the following.   **[4×4=16]**

   a)   Explain the advantages and disadvantages of Apriori algorithm.

   b)   Explain four data types in R.

   c)   Explain four types of Correlation.

   d)   Explain Statistical Inference with suitable diagram.

   e)   Differentiate between structured and Unstructured Data.

P.T.O.

Q3) Attempt any FOUR of the following : [4×4=16]

    a) Explain 5V's of Big Data.

    b) Explain the phases of Data Analytics Life Cycle.

    c) Explain *for* loop in R programming with syntax and example.

    d) Explain Naive Bayes Algorithm in detail.

    e) Explain Head ( ) and Tail ( ) functions in *dplyr* package.

Q4) Attempt any FOUR of the following : [4×4=16]

    a) Explain four applications of Big Data.

    b) What is population? Explain different types of population.

    c) Write an R program to sort a Vector in ascending and descending order.

    d) Write an R program to create a simple bar plot of five subject's marks.

    e) Write an R program to calculate Multiplication Table.

Q5) Write a short note on any TWO of the following : [2×3=6]

    a) Advantages of SVM algorithm.

    b) Statistical Modeling.

    c) Data Analysis.

☩ ☩ ☩