ST. JOSEPH'S COLLEGE (AUTONOMOUS), BANGALORE – 27
OPEN ELECTIVE – II SEMESTER
END SEMESTER EXAMINATION – APRIL 2022
(EXAMINATION CONDUCTED IN JULY-AUGUST 2022)
DAOE2: MACHINE LEARNING USING R PROGRAMMING

Time: 2 hrs	**Maximum marks: 60 marks**

This question paper contains SEVEN printed pages and one part

PART A

Answer all the questions   **60 x 1 = 60**

1. Give the function for creating vector in R.
   a. a(2,8,6,9)
   b. a = c(3,5,6,2)
   c. c = (1,2,8,6)
   d. a = 9,5,4,3

2. A _____ is a two-dimensional rectangular data set.
   a. Vector
   b. Lists
   c. Matrix
   d. Functions

3. What is the mean of the following numbers: 14, 24, 87, 40, 50?
   a. 43
   b. 34
   c. 56
   d. None of the above

4. _____ is not a measure of central tendency.
   a. Mode
   b. Mean
   c. Range
   d. Median

5. Mode refers to the value within a series that occurs _____ number of times.
   a. Maximum
   b. Minimum
   c. Zero
   d. Infinite

6. _____ divides the data into four equal parts.
   a. Median
   b. Quartiles
   c. Mean
   d. None of the above

7. From the following data, what is the value of the median?
         15 27 37 28 23 30 31 40 35
   a. 27
   b. 23
   c. 25.22
   d. 30

8. A _____ is an R-object which can contain many different types of elements inside it.
   a. Vector
   b. Lists
   c. Matrix
   d. Functions

9. Which function takes a dim attribute which creates the required number of dimensions?
   a. Vector
   b. Array
   c. Matrix
   d. Lists

10. The _____ function creates a regular sequence of values to form a vector.
   a. sequel
   b. Rep
   c. seq
   d. Grep

11. _____ are similar to matrices but can have more than two dimensions.
   a. Functions
   b. Packages
   c. Arrays
   d. Columns

12. . Find the following type of vector?

a = c(1,2,5.3,6,-2,4)

   a. Numeric
   b. Character
   c. Integer
   d. Logical

13. An ordered collection of objects or components are called _____
   a. Data frames
   b. Datasets
   c. Databases
   d. Lists

14. The function that returns the smallest integer which is larger than or equal to x is
_____.
   a. trunc()
   b. floor()
   c. ceiling()
   d. abs()
15. The function that converts lowercase to uppercase string is _____.
   a. tolower()
   b. tomax()
   c. tomin()
   d. toupper()
16. The function that will give the maximum value in a given set of value is _____.
   a. min()
   b. large()
   c. max()
   d. maximum()
17. The function that will give the sum value in a given set of value is _____.
   a. sum()
   b. add()
   c. max()
   d. maximum()
18. The function that will give the standard deviation in a given set of value is _____.
   a. dev()
   b. sd()
   c. sdev()
   d. maximum()
19. The function that will print the last 6 rows of a dataset is _____.
   a. tail()
   b. trail()
   c. head()
   d. header()
20. The function that splits the given string is _____.
   a. split()
   b. strsplit()
   c. stringsplit()
   d. splitstring()
21. The function that will print the first 6 rows of a dataset is _____.
   a. tail()
   b. trail()
   c. head()
   d. header()
22. The function that returns the smallest integer which is larger than or equal to x is
_____.
   a. trunc()
   b. floor()
   c. ceiling()
   d. abs()
23. What is Machine learning?
   a. The autonomous acquisition of knowledge through the use of computer programs
   b. The autonomous acquisition of knowledge through the use of manual programs
   c. The selective acquisition of knowledge through the use of computer programs
   d. The selective acquisition of knowledge through the use of manual programs

24. ML is a field of AI consisting of learning algorithms that?
    a. Improve their performance
    b. At executing some task
    c. Over time with experience
    d. All of the above
25. Which of the following are ML methods?
    a. based on human supervision
    b. supervised Learning
    c. semi-reinforcement Learning
    d. All of the above
26. Machine learning algorithms build a model based on sample data, known as ……..
    a. Training Data
    b. Transfer Data
    c. Data Training
    d. None of the above
27. Machine learning is a subset of …………….
    a. Deep Learning
    b. Artificial Intelligence
    c. Data Learning
    d. None of the above
28. Machine learning is an application of _____.
    a. Blockchain
    b. Artificial Intelligence
    c. Both A and B
    d. None of the above
29. _____ is the machine learning algorithms that can be used with labeled data.
    a. Regression algorithms
    b. Clustering algorithms
    c. Association algorithms
    d. All of the above
30. Which Language is Best for Machine Learning?
    a. C
    b. Java
    c. Python
    d. HTML
31. Automatic Speech Recognition systems find a wide variety of applications in the _____ domains.
    a. Medical Assistance
    b. Industrial Robotics
    c. Defence & Aviation
    d. All of the above
32. Which machine learning models are trained to make a series of decisions based on the rewards and feedback they receive for their actions?
    a. Supervised learning
    b. Unsupervised learning
    c. Reinforcement learning
    d. All of the above
33. _____ is simplest class of analytics.
    a. Descriptive
    b. Predictive
    c. Prescriptive

d. Summarization

34. Which of the following involves predicting a categorical response?
    a. Regression
    b. Summarization
    c. Clustering
    d. Classification

35. Which of the following simple the square root of σ square?
    a. Variance
    b. Standard deviation
    c. Mean
    d. Median

36. Every _____ function has a probability distribution function.
    a. Continuous
    b. Discrete
    c. Categorical
    d. Random

37. Which distribution looks like a norm distribution but with very heavy tails?
    a. Simple
    b. Discrete
    c. Continuous
    d. Cauchy

38. The _____ is represented by the location parameter.
    a. Median
    b. Mode
    c. Mean
    d. Variance

39. _____ are commonly introduced as representations of quantities that have magnitude and direction.
    a. Lists
    b. Data sets
    c. Data frames
    d. Vectors

40. _____ involves predicting a response with meaningful magnitude, such as quantity sold, stock price, or return on investment.
    a. Regression
    b. Summarization
    c. Clustering
    d. Classification

41. Which of the following is preferred for text analytics?
    a. R
    b. Python
    c. S
    d. EDR

42. _____ is simplest class of analytics.
    a. Descriptive
    b. Predictive
    c. Prescriptive
    d. Summarization

43. _____ phase of the data analytics lifecycle usually takes the longest time.
    a. Phase 2: Data Preparation

b. Phase 3: Model Planning
c. Phase 4: Model Building
d. Phase 5: Communicate Results

44. When data are collected in a statistical study for only a portion or subset of all elements of interest we are using _____ .
   a. Sample
   b. Parameter
   c. Population
   d. None

45. The dependent variable is also called _____ .
   a. Regression
   b. Regressand
   c. Continuous variable
   d. Independent

46. The dependent variable is also called _____ .
   a. Regressor
   b. Predictand variable
   c. Explained variable
   d. All of these

47. What will be the output of the following statement?
   x = -10
   print(abs(x)
   a. 3
   b. 6
   c. 8
   d. 10

48. All columns in a matrix must have the same mode and the _____ length.
   a. Different
   b. Same
   c. May be the same
   d. May be different

49. Which function takes a dim attribute which creates the required number of dimensions?
   a. Vector
   b. Array
   c. Matrix
   d. Lists

50. By which function we can create data frames
   a. Data.frames()
   b. Data.sets ()
   c. Function ()
   d. C ()

51. Which of the following step is performed by data scientist after acquiring the data?
   a. Data Cleansing
   b. Data Integration
   c. Data replication
   d. All the mentioned

52. Which of the following is one of the key data science skills?
   a. Statistics

b. Machine Learning
c. Data Visualization
d. All of the mentioned

53. Logistic regression is for predicting_____
    a. Continuous variable
    b. Categorical Variable
    c. continuous and categorical
    d. None

54. _____is not a data analytics activity.
    a. Case study
    b. data reporting
    c. Data cleaning
    d. None

55. Which of the following is considered data science?
    a. Case study
    b. Qualitative analysis
    c. Qualitative analysis
    d. Forecasting

56. Data science is the process of diverse set of data through_____
    a. organizing data
    b. processing data
    c. analysing data
    d. All of the mentioned

57. R script files has an extension_____
    a. .RP
    b. .R
    c. .Rn
    d. None

58. Which of the following is related to the meaning of the term analytics?
    a. Analytics about explaining past events
    b. Analytics about explaining present events
    c. Analytics about predicting future events
    d. None

59. R language has superficial similarity with _____
    a. C
    b. Python
    c. MATLAB
    d. SAS

60. _____ function is used to watch for all available packages in library.
    a. lib()
    b. fun.lib()
    c. libr()
    d. library()