

# Online Examinations (Even Sem/Part-I/Part-II Examinations 2020 - 2021)

Course Name - Introduction to Data Sciences

Course Code - BCS603B

\* You can submit the form ONLY ONCE.

\* Fill the following information for further process.

\* Required

1. Email \*

---

2. Name of the Student \*

---

3. Enter Full Student Code \*

---

4. Enter Roll No \*

---

5. Enter Registration No \*

---

6. Enter Course Code \*

---

7. Enter Course Name \*

---

8. \*

*Mark only one oval.*

- Diploma in Pharmacy
- Bachelor of Pharmacy
- B.TECH.(CSE)
- B.TECH.(ECE)
- BCA
- B.SC.(CS)
- B.SC.(BT)
- B.SC.(ANCS)
- B.SC.(HN)
- B.Sc.(MM)
- B.A.(MW)
- BBA
- [B.COM](#)
- B.A.(JMC)
- BBA(HM)
- BBA(LLB)
- B.OPTOMETRY
- B.SC.(MB)
- B.SC.(MLT)
- B.SC.(MRIT)
- B.SC.(PA)
- LLB
- [B.SC\(IT\)-AI](#)
- B.SC.(MSJ)
- Bachelor of Physiotherapy
- B.SC.(AM)
- Dip.CSE
- Dip.ECE
- [DIP.EE](#)
- DIP.CE

- [DIP.ME](#)
- PGDHM
- MBA
- M.SC.(BT)
- M.TECH(CSE)
- LLM
- M.A.(JMC)
- M.A.(ENG)
- M.SC.(MATH)
- M.SC.(MB)
- MCA
- M.SC.(MSJ)
- M.SC.(AM)
- M.SC.CS)
- M.SC.(ANCS)
- M.SC.(MM)
- B.A.(Eng)

Answer all the questions. Each question carry one mark.

9. 1. Classification accuracy is

*Mark only one oval.*

- A subdivision of a set of examples into a number of classes
- Measure of the accuracy, of the classification of a concept that is given by a certain theory
- The task of assigning a classification to a set of examples
- None of these

10. 2. KDD (Knowledge Discovery in Databases) is referred to

*Mark only one oval.*

- Non-trivial extraction of implicit previously unknown and potentially useful information from data
- Set of columns in a database table that can be used to identify each record within this table uniquely.
- collection of interesting and useful patterns in a database
- None of these

11. 3. Which of the following method options is provided by train function for bagging?

*Mark only one oval.*

- bagEarth
- treebag
- bagFDA
- all of these

12. 4. How calculations work in TensorFlow

*Mark only one oval.*

- Through vector multiplications
- Through RDDs
- Through Computational Graphs
- Through map reduce tasks

13. 5. Branch of statistics which deals with findings of solution in the field of medicine, education and economics is classified as

*Mark only one oval.*

- economic statistics
- applied statistics
- mathematical statistics
- industry statistics

14. 6. Algorithm is

*Mark only one oval.*

- It uses machine-learning techniques. Here program can learn from past experience and adapt themselves to new situations
- Computational procedure that takes some value as input and produces some value as output
- Science of making machines performs tasks that would require intelligence when performed by humans
- None of these

15. 7. Hidden knowledge referred to

*Mark only one oval.*

- A set of databases from different vendors, possibly using different database paradigms.
- An approach to a problem that isn't guaranteed to work but performs well in most cases
- Information that is hidden in a database and that can't be recovered by a simple SQL query.
- None of these

16. 8. Point out the correct statement.

*Mark only one oval.*

- Combining classifiers improves interpret ability
- Combining classifiers reduces accuracy
- Combining classifiers improves accuracy
- All of these

17. 9. Which of the following can be used to create the most common graph types?

*Mark only one oval.*

- qplot
- quickplot
- plot
- All of these

18. 10. Which of the following measures of central tendency will always change if a single value in the data changes?

*Mark only one oval.*

- Mean
- Median
- Mode
- All of these

19. 11. Data selection is

*Mark only one oval.*

- The actual discovery phase of a knowledge discovery process
- The stage of selecting the right data for a KDD process
- A subject-oriented integrated time variant non-volatile collection of data in support of management
- None of these

20. 12. R functionality is divided into a number of \_\_\_\_\_

*Mark only one oval.*

- Packages
- Functions
- Domains
- None of these

21. 13. Point out the wrong combination.

*Mark only one oval.*

- True negative=correctly rejected
- False negative=correctly rejected
- False positive=correctly identified
- All of these

22. 14. Attributes of an object (if any) can be accessed using the \_\_\_\_\_ function.

*Mark only one oval.*

- objects()
- attrib()
- attributes()
- obj()

23. 15. Figure out the wrong statement

*Mark only one oval.*

- k-means clustering is a method of vector quantization
- k-means clustering aims to partition n observations into k clusters
- k-nearest neighbor is same as k-means
- none of these

24. 16. \_\_\_\_\_ is used to make predictions about unknown future events?

*Mark only one oval.*

- Descriptive analysis
- Predictive analysis
- Descriptive and Predictive analysis
- None of these

25. 17. Which tool is best suited for solving Deep Learning problems

*Mark only one oval.*

- R
- Sk-learn
- Excel
- TensorFlow

26. 18. Which of the following can be considered as a random variable?

*Mark only one oval.*

- The outcome from the roll of a die
- The outcome of flip of a coin
- The outcome of exam
- All of these

27. 19. Which of the following techniques can be used for normalization in text mining?

*Mark only one oval.*

- Stemming
- Lemmatization
- Stop Word Removal
- Stemming and Lemmatization

28. 20. \_\_\_\_\_ initiates an infinite loop right from the start.

*Mark only one oval.*

- Never
- Repeat
- Break
- Set

29. 21. R was named partly after the first names of \_\_\_\_ R authors?

*Mark only one oval.*

- 2
- 1
- 3
- 4

30. 22. Point out the wrong statement

*Mark only one oval.*

- The trapezoidal rule is used to compute the area under the ROC curve
- For regression, the relationship between each predictor and the outcome is evaluated
- An argument, para, is used to pick the model fitting technique
- All of these

31. 23. Point out the wrong statement:

*Mark only one oval.*

- Randomized studies are not used to identify causation
- Complication approached exist for inferring causation
- Causal relationships may not apply to every individual
- All of these

32. 24. When performing regression or classification, which of the following is the correct way to preprocess the data?

*Mark only one oval.*

- Normalize the data → PCA → training
- PCA → normalize PCA output → training
- Normalize the data → PCA → normalize PCA output → training
- None of these

33. 25. How many types of atomic vectors are present?

*Mark only one oval.*

- 3
- 4
- 5
- 6

34. 26. A definition of a concept is-----if it recognizes all the instances of that concept

*Mark only one oval.*

- Complete
- Consistent
- Constant
- None of these

35. 27. Inductive logic programming is

*Mark only one oval.*

- A class of learning algorithms that try to derive a Prolog program from examples
- A table with n independent attributes can be seen as an n-dimensional space
- Prediction made using an extremely simple method, such as always predicting the same output
- None of these

36. 28. Find out the correct statement

*Mark only one oval.*

- Prediction with regression is easy to implement
- Prediction with regression is easy to interpret
- Prediction with regression performs well when linear model is correct
- All of these

37. 29. To find the minimum or the maximum of a function, we set the gradient to zero because

*Mark only one oval.*

- The value of the gradient at extrema of a function is always zero
- Depends on the type of problem
- All of these
- None of these

38. 30. What are the four main things we should know before studying data analysis?

*Mark only one oval.*

- Descriptive statistics
- Inferential statistics
- Distributions (normal distribution / sampling distribution)
- Hypothesis testing

39. 31. Background knowledge referred to

*Mark only one oval.*

- Additional acquaintance used by a learning algorithm to facilitate the learning process
- A neural network that makes use of a hidden layer
- It is a form of automatic learning.
- None of these

40. 32. Heterogeneous databases referred to

*Mark only one oval.*

- A set of databases from different vendors, possibly using different database paradigms
- An approach to a problem that is not guaranteed to work but performs well in most cases.
- Information that is hidden in a database and that cannot be recovered by a simple SQL query
- None of these

41. 33. Which of the following is used to assist the quantitative trader in the development?

*Mark only one oval.*

- quantmod
- quantile
- quantity
- mboost

42. 34. Which of the following is required by K-means clustering?

*Mark only one oval.*

- defined distance metric
- number of clusters
- initial guess as to cluster centroids
- all of these

43. 35. What happens to the confidence interval when we introduce some outliers to the data?

*Mark only one oval.*

- Confidence interval is robust to outliers
- Confidence interval will increase with the introduction of outliers.
- Confidence interval will decrease with the introduction of outliers
- We cannot determine the confidence interval in this case

44. 36. Adaptive system management is

*Mark only one oval.*

- It uses machine -learning techniques. Here program can learn from past experience and adapt themselves to new situations
- Computational procedure that takes some value as input and produces some value as output.
- Science of making machines performs tasks that would require intelligence when performed by humans
- None of these

45. 37. Classification task referred to

*Mark only one oval.*

- A subdivision of a set of examples into a number of classes
- A measure of the accuracy, of the classification of a concept that is given by a certain theory
- The task of assigning a classification to a set of examples
- None of these

46. 38. Which of the following is correct about regularized regression?

*Mark only one oval.*

- Cannot help with model selection
- Can help with bias trade-off
- Cannot help with variance trade-off
- All of these

47. 39. Which of the following is not a machine learning algorithm?

*Mark only one oval.*

- SVM
- Random forest are easy to interpret but often very accurate
- SVG
- None of these

48. 40. Which of the following can be considered as an object attribute?

*Mark only one oval.*

- Dimensions
- class
- length
- All of these

49. 41. How many steps does the predictive analysis process contain?

*Mark only one oval.*

8

7

5

6

50. 42. \_\_\_\_ programming language is a dialect of S.

*Mark only one oval.*

B

C

D

R

51. 43. Which of the following is also referred to as a random variable?

*Mark only one oval.*

stochast

aleatory

eliette

All of these

52. 44. Which of the following is a reasonable way to select the number of principal components "k"?

*Mark only one oval.*

- Choose k to be the smallest value so that at least 99% of the variance is retained.
- Choose k to be 99% of m ( $k = 0.99 * m$ , rounded to the nearest integer).
- Choose k to be the largest value so that 99% of the variance is retained
- Use the elbow method

53. 45. The most widely used metrics and tools to assess a classification model are:

*Mark only one oval.*

- Confusion matrix
- Cost-sensitive accuracy
- Area under the ROC curve
- All of these

54. 46. R is an interpreted language so it can access through\_\_\_\_\_?

*Mark only one oval.*

- Disk operating system
- User interface operating system
- Operating system
- Command line interpreter

55. 47. Which of the following functions tracks the changes in model statistics?

*Mark only one oval.*

- varImp
- varImpTrack
- findTrack
- None of these

56. 48. Which of the following is commonly referred to as 'data fishing'?

*Mark only one oval.*

- Data bagging
- Data booting
- Data merging
- None of these

57. 49. What is `pca.components_` in Sklearn?

*Mark only one oval.*

- Set of all eigen vectors for the projection space
- Matrix of principal components
- Result of the multiplication matrix
- None of these

58. 50. \_\_\_\_\_ and \_\_\_\_\_ are types of matrices functions?

*Mark only one oval.*

- Apply and supply
- Apply and lapply
- Both of these
- None of these

59. 51. Data mining is

*Mark only one oval.*

- The stage of selecting the right data for a KDD process.
- A subject-oriented integrated time variant non-volatile collection of data in support of management.
- The actual discovery phase of a knowledge discovery process.
- None of these

60. 52. Prediction is

*Mark only one oval.*

- The result of the application of a theory or a rule in a specific case
- One of several possible enters within a database table that is chosen by the designer as the primary means of accessing the data in the table.
- Discipline in statistics that studies ways to find the most interesting projections of multi-dimensional spaces
- None of these

61. 53. Which of the following is one of the largest boost subclass in boosting?

*Mark only one oval.*

- variance boosting
- gradient boosting
- mean boosting
- all of these

62. 54. Which of the following is a disadvantage of decision trees?

*Mark only one oval.*

- Factor analysis
- Decision trees are robust to outliers
- Decision trees are prone to be over fit
- None of these

63. 55. Which function is used to create the vector with more than one element?

*Mark only one oval.*

- Library()
- plot()
- c()
- par()

64. 56. Classification is

*Mark only one oval.*

- A measure of the accuracy of the classification of a concept that is given by a certain theory
- A subdivision of a set of examples into a number of classes.
- The task of assigning a classification to a set of examples.
- None of these

65. 57. Heuristic is

*Mark only one oval.*

- A set of databases from different vendors, possibly using different database paradigms
- An approach to a problem that is not guaranteed to work but performs well in most cases
- Information that is hidden in a database and that cannot be recovered by a simple SQL query.
- None of these

66. 58. Predicting with trees evaluate \_\_\_\_\_ within each group of data.

*Mark only one oval.*

- equality
- homogeneity
- heterogeneity
- all of these

67. 59. Which of the following clustering requires a merging approach?

*Mark only one oval.*

- Partitional
- Hierarchical
- Naive Bayes
- None of these

68. 60. Tools such decision making by nominal groups, brainstorming and term buildings are all considered as

*Mark only one oval.*

- serial tools
- behavioral tools
- statistical tools
- parallel tools

69. 61. Bayesian classifiers is

*Mark only one oval.*

- A class of learning algorithm that tries to find an optimum classification of a set of examples using the probabilistic theory.
- Any mechanism employed by a learning system to constrain the search space of a hypothesis
- An approach to the design of learning algorithms that is inspired by the fact that when people encounter new situations, they often explain them by reference to familiar experiences, adapting the explanations to fit the new situation.
- None of these

70. 62. Discovery is

*Mark only one oval.*

- It is hidden within a database and can only be recovered if one is given certain clues (an example IS encrypted information).
- The process of executing implicit previously unknown and potentially useful information from data
- An extremely complex molecule that occurs in human chromosomes and that carries genetic information in the form of genes.
- None of these

71. 63. Which of the following methods are present in caret for regularized regression?

*Mark only one oval.*

- ridge
- lasso
- relaxo
- all of these

72. 64. Which of the following is a categorical outcome?

*Mark only one oval.*

- RMSE
- RSquared
- Accuracy
- All of these

73. 65. Five numbers are given: (5, 10, 15, 5, 15). Now, what would be the sum of deviations of individual data points from their mean?

*Mark only one oval.*

- 10
- 25
- 50
- 0

74. 66. How many data types are present in R?

*Mark only one oval.*

- 5
- 4
- 6
- 7

75. 67. Finally, in \_\_\_\_\_ R version 1.0.0 was released to the public.

*Mark only one oval.*

- 2005
- 2000
- 2012
- 2010

76. 68. Which of the following functions is associated with a continuous random variable?

*Mark only one oval.*

- pdf
- pmv
- pmf
- All of these

77. 69. What is a sentence parser typically used for?

*Mark only one oval.*

- It is used to parse sentences to check if they are utf-8 compliant
- It is used to parse sentences to derive their most likely syntax tree structures.
- It is used to parse sentences to assign POS tags to all tokens.
- It is used to check if sentences can be parsed into meaningful tokens.

78. 70. Which of the following is also referred to as random variable?

*Mark only one oval.*

- stochast
- aleatory
- eliette
- All of the above

---

This content is neither created nor endorsed by Google.

Google Forms