



# BRAINWARE UNIVERSITY

Term End Examination 2023-2024  
Programme – MCA-2022  
Course Name – Applied Data Science  
Course Code - MCA401B  
( Semester IV )

Full Marks : 60

Time : 2:30 Hours

[The figure in the margin indicates full marks. Candidates are required to give their answers in their own words as far as practicable.]

## Group-A

(Multiple Choice Type Question)

1 x 15=15

1. Choose the correct alternative from the following :

- (i) Predictive analysis is defined as branch of \_\_\_\_\_ analysis.
- |                  |                  |
|------------------|------------------|
| a) Advanced      | b) Core          |
| c) Both of these | d) None of these |
- (ii) Identify the number of steps the predictive analysis process contains.
- |      |      |
|------|------|
| a) 5 | b) 6 |
| c) 7 | d) 8 |
- (iii) \_\_\_\_\_ is identified to make predictions about unknown future events.
- |                        |                         |
|------------------------|-------------------------|
| a) Predictive analysis | b) Descriptive analysis |
| c) Both of these       | d) None of these        |
- (iv) Bayesian classifiers is defined as
- |  |  |
|--|--|
| a) A class of learning algorithm that tries to find an optimum classification of a set of examples using the probabilistic theory.   | b) Any mechanism employed by a learning system to constrain the search space of a hypothesis |
| c) 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 | d) None of these   |
- (v) Who developed R?
- |                      |                    |
|----------------------|--------------------|
| a) Dennis Ritchie    | b) John Chambersii |
| c) Bjarne Stroustrup | d) None of these   |
- (vi) Euclidean distance measure is calculated in
- |  |   |
|--|---|
| a) A stage of the KDD process in which new data is added to the existing selection | b) The process of finding a solution for a problem simply by enumerating all possible |
|--|---|

- solutions according to some pre-defined order and then testing them
- c) The distance between two points as calculated using the Pythagoras theorem
- (vii) Hidden knowledge relates to
- a) A set of databases from different vendors, possibly using different database paradigms
- b) An approach to a problem that is not guaranteed to work but performs well in most cases
- c) Information that is hidden in a database and that cannot be recovered by a simple SQL query
- d) None of these
- (viii) Enumeration is related to
- a) A stage of the KDD process in which new data is added to the existing selection
- b) The process of finding a solution for a problem simply by enumerating all possible solutions according to some pre-defined order and then testing them
- c) The distance between two points as calculated using the Pythagoras theorem
- d) None of these
- (ix) KDD is interpreted as (Knowledge Discovery in Databases)
- a) Non-trivial extraction of implicit previously unknown and potentially useful information from data
- b) Set of columns in a database table that can be used to identify each record within this table uniquely
- c) collection of interesting and useful patterns in a database
- d) None of these
- (x) What is Naive prediction ?
- a) A class of learning algorithms that try to derive a Prolog program from example
- b) A table with n independent attributes can be seen as an n- dimensional space
- c) A prediction made using an extremely simple method, such as always predicting the same output
- d) None of these
- (xi) Determine the wrong statement
- a) The trapezoidal rule is used to compute the area under the ROC curve
- b) For regression, the relationship between each predictor and the outcome is evaluated
- c) An argument, para, is used to pick the model fitting technique
- d) All of these
- (xii) To evaluate the minimum or the maximum of a function, we set the gradient to zero because
- a) The value of the gradient at extrema of a function is always zero
- b) Depends on the type of problem
- c) Both The value of the gradient at extrema of a function is always zero and Depends on the type of problem
- d) None of these
- (xiii) What is the purpose of performing cross-validation?
- a) To assess the predictive performance of the models
- b) To judge how the trained model performs outside the sample on test data
- c) To assess the predictive performance of the models and To judge how the trained model performs outside the sample on test data
- d) None of these
- (xiv) Which of the following is distinguished as example of feature extraction?
- a) Constructing bag of words vector from an email
- b) Applying PCA projects to a large high-dimensional data

