



BRAINWARE UNIVERSITY

Term End Examination 2022

Programme – B.Tech.(CSE)-2017/B.Tech.(CSE)-2018/B.Tech.(CSE)-2019

Course Name – Data Analytics

Course Code - BCSE701/PEC-702A

(Semester VII)

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) A fair six-sided die is rolled twice. Calculate What is the probability of getting 2 on the first roll and not getting 4 on the second roll?
 - a) $1/36$
 - b) $1/18$
 - c) $5/36$
 - d) $1/6$
- (ii) Suppose you have trained a logistic regression classifier and it outputs a new example x with a prediction $h_0(x) = 0.2$. This determine
 - a) our estimate for $P(y=1 | x)$
 - b) our estimate for $P(y=0 | x)$
 - c) All of these
 - d) None
- (iii) For t distribution, increasing the sample size, the effect will be apply on
 - a) degrees of freedom
 - b) the t -ratio
 - c) standard error of the means
 - d) all of these
- (iv) Suppose you are using a bagging based algorithm say a random forest in model building. Select which of the following can be true? 1.Number of tree should be as large as possible 2.You will have interpretability after using random forest
 - a) 1
 - b) 2
 - c) 1 and 2
 - d) none of these
- (v) Categorize, The process of adjusting the weight is known as
 - a) activation
 - b) synchronization
 - c) learning
 - d) none of the mentioned
- (vi) Select Which of the following mentioned standard probability density functions is applicable to discrete random variables?
 - a) gaussian distribution
 - b) poisson distribution
 - c) rayleigh distribution
 - d) exponential distribution
- (vii) The denominator (bottom) of the z -score formula is defined as
 - a) the standard deviation
 - b) b.the difference between a score and the mean
 - c) the range
 - d) the mean
- (viii) select which of the following is the advantage/s of decision trees?

OR

Analyze why SVMs are often more accurate than logistic regression with examples. (5)
12. Assess what are the best practices in big data analytics. (5)

OR

Evaluate the techniques used in big data analytics. (5)
