

Bihar Engineering University, Patna
End Semester Examination - 2022

Course: B.Tech.
Code: 105701

Semester: 7
Subject: Data Science

Time: 03 Hours
Full Marks: 70

Instructions:-

- (i) The marks are indicated in the right-hand margin.
- (ii) There are NINE questions in this paper.
- (iii) Attempt FIVE questions in all.
- (iv) Question No. 1 is compulsory.

Q.1 Answer in brief any seven questions:

[2 x 7 = 14]

- (a) Define Data Science.
- (b) Point out the rules for dot product of two vectors.
- (c) What is Machine Learning?
- (d) What is the reason for the word "Naïve" in Naïve Bayes classification?
- (e) Define data visualization in machine learning.
- (f) Point out the component of Data Science.
- (g) What is Object Recognition?
- (h) Compare Big data with Data Science.
- (i) Differentiate Data Mining and Data Science.
- (j) Compare various data science languages.

Q.2 (a) What is Bigdata?

[4]

(b) Describe the main features of a big data in detail.

[10]

Q.3 Describe life cycle of Data Science with neat diagram.

[14]

Q.4 Explain the following in detail.

(a) Map Reduce

[7]

(b) YARN

[7]

Q.5 Describe vectors and various operations on vectors with routines, example and diagram.

[14]

Q.6 Develop a routine to demonstrate Binomial Distribution and Normal Distribution.

[14]

Q.7 Explain the following with suitable example.

(a) Simple Linear Regression

[7]

(b) Multiple Regression

[7]

Q.8 (a) Describe numpy in detail.

[7]

(b) Write a python program that uses numpy and explain it.

[7]

Q.9 Discuss various sub-processes involved in the complete process of data science for weather prediction.

[14]

