|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |

**E Reg. No. :**

**Question Paper Code: 52P31**

 M.E DEGREE EXAMINATION, NOV 2017

Second Semester

Computer Science and Engineering

15PCS201 - DATA SCIENCE AND BIG DATA ANALYTICS

(Regulation 2015)

Duration: Three hours Maximum: 100 Marks

Answer ALL Questions

PART - A (5 x 20 = 100 Marks)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1. | (a) | Describe the Challenges of the current analytical Architecture for Data Scientists. | CO1- U |  (20) |
|  |  | Or |  |  |
|  | (b) | Explain the concept of an analytic sand box.  | CO1- U |  (20) |
|  |  |  |  |  |
| 2. | (a) | Illustrate the discovery phase for the business model  | CO2- U |  (20) |
|  |  | Or |  |  |
|  | (b) | Describe about Initial Hypothesis testing  | CO2- U | (20) |
|  |  |  |  |  |
| 3. | (a) | R programming language has several packages for data science which are meant to solve a specific problem, how do you decide which one to use? | CO3-App | (20) |
|  |  | Or |  |  |
|  | (b) | Show the statistics for Household Income and visualize data with “R”. | CO3-App | (20) |
|  |  |  |  |  |
| 4. | (a) | Illustrate Linear regression with example.  | CO4- U | (20) |
|  |  | Or |  |  |
|  | (b) | Explain the classification techniques with example. | CO4- U | (20) |
|  |  |  |  |  |
| 5. | (a) | Explain the complexity theory for Map-Reduce? What is reducer size and replication rate  | CO5- U | (20) |
|  |  | Or |  |  |
|  | (b) | Explain MADlib functions with example. | CO5- U | (20) |
|  |  |  |  |  |