Reg. No. : ☐☐☐☐☐☐☐☐☐☐

Question Paper Code: 52231

M.E. DEGREE EXAMINATION, JUNE 2016

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) Discuss Bigdata in terms of three dimensions volume, variety, and velocity.   (20)

Or

  (b) Discuss industry examples of Bigdata in detail.   (20)

2. (a) Briefly explain the steps in the life cycle approach.   (20)

Or

  (b) Discuss the key roles for successful analytic project management.   (20)

3. (a) Discuss how *R* will be a well-defined integrated suite of software for data
      manipulation, calculation and graphical display.   (20)

Or

  (b) What are statistical models? Discuss the evaluation methodologies.   (20)

4. (a) Discuss about the classifier based on Bayes theorem with example. (20)

Or

(b) Explain the linear and logistic regression model; also discuss which is better and when? (20)

5. (a) Discuss YARN (map-reduce 2) in detail with failures in classic map-reduce. (20)

Or

(b) Elaborate on HivQL data manipulation queries in detail. (20)

———————————